

AMERICAN JOURNAL of PHYSICS

A Journal Devoted to the Instructional and Cultural Aspects of Physical Science

VOLUME 15, NUMBER 5

SEPTEMBER-OCTOBER 1947

The Franck-Condon Principle and Related Topics*

E. U. CONDON

National Bureau of Standards, Washington 25, District of Columbia

A REVIEW of the historical development and present status of certain topics in molecular physics may serve as a reminder of the progress made in the past two decades. Let us recall that in 1925 there was no subject of nuclear physics, and high voltage equipment was almost non-existent except in the x-ray laboratories. About all that was known about cosmic rays was that the ionization increases in high balloon flights in the atmosphere and that its cause penetrates deeply into lakes. Nobody had heard of closed electron bands in a solid, and the word semiconductor was essentially unknown.

"What did physicists find to be interested in?" young physicists may wonder. The answer is, of course, that they worried about a lot of things that are nowadays taught to and learned by beginning graduate students with such glibness that it is hard to realize there ever was a time when these things were not known.

In 1925 there was no quantum mechanics. Though de Broglie's thesis was published in 1924, I have never met anyone who read it seriously until later. We worked with quantization by the $\int p dq = nh$ method of Bohr and Sommerfeld and tried to get at radiative transition probabilities

by approximate and unclearly formulated procedure based on the Bohr correspondence principle, according to which quantum radiative jumps were associated with, or set in correspondence with certain terms in the Fourier analysis of the quasiperiodic motions in the mechanical system.¹

The Bohr theory had given a beautiful account of the spectrum of atomic hydrogen and of the arc spectra of the alkalis. However, why the D lines of sodium were double was a great mystery that was only cleared up by the electron-spin hypothesis of Goudsmit and Uhlenbeck. Similarly, the old quantum theory had thrown a good deal of light on the infra-red and electronic spectra of diatomic molecules. Pure rotation, rotation-vibration, and electronic band systems were recognized, analyzed and utilized to get interesting quantitative data on important molecules. Nevertheless, here too there were puzzling things; for example, the isotope shift between HCl^{35} and HCl^{37} , and in other cases, required that the vibrational levels be assigned half-integral quantum numbers.

First Ideas on the Principle

At that time I was a graduate student in the University of California in Berkeley. Ernest Lawrence was a graduate student at Yale, and

* Address of the Retiring President, American Physical Society, New York, January 31, 1947. My apology for including my own name in the title of this paper is that it conforms to a well-established usage. In the text I shall refer simply to the Principle. Professor Rabi once said to me that the Principle was a great boon to lecturers in courses on atomic physics, for it is so easy to understand that they are always assured of one or two lectures which they do not have to prepare.

¹ W. Lenz, *Zeits. f. Physik* 25, 299 (1924). This paper represents an attempt to deal with the problem of nuclear transitions associated with electronic jumps by means of the correspondence principle.

there were no cyclotrons in the world—not one, not even in Berkeley. Professors R. T. Birge and L. B. Loeb were the stimulating research leaders of the department, the former on molecular spectra, the latter on every phase of electric conduction in gases.

During the year 1925–26, Birge conducted a seminar on molecular spectra which it is impossible for me to praise too highly. Then as always his work was distinguished by the most exact and painstaking scrutiny of the data in their relation to the theory. I still remember how exciting it was when he first showed, with precision, that the swelling of molecules by rotation, as inferred from their pure rotation spectrum, agreed with the values found from the vibration-rotation spectrum and from electronic systems. It was he who had carefully compiled all the existing data on analyzed electronic band systems—then about a dozen in number—and who recognized the basic empirical facts of the intensity distribution, later to be explained by the Principle. I mention this point so explicitly because without his stimulating guidance I would never have been aware of the problem of intensity distribution in band systems.

In Göttingen, Professor James Franck was very much interested in photochemical reactions and, in particular, the dissociation of iodine vapor by absorption of light. He gave a paper before the Faraday Society in London² in which the basic idea of the Principle was first presented in connection with this problem.

Proof sheets of this paper were sent by Franck to his student, Dr. Hertha Sponer, who was in Berkeley that year on an International Education Board Fellowship. She let me read them, and, because I had just learned the empirical problem about intensity distribution in band sys-

tems from Birge's seminar, it was immediately evident how to generalize Franck's ideas a little more in order to get the full story. What is more, from Birge's seminar I had at hand a good critical compilation of the existing data which made possible a quick quantitative test of the ideas.

This work was all done in a few days. Doctor Sponer showed me Franck's paper one afternoon, and a week later all the quantitative work for my 1926 paper³ was done. But let us see just what was the situation.

Figure 1 is from Franck's paper.² He pointed out that, owing to the large masses of the nuclei in a molecule, their relative momentum cannot be directly affected by an electronic transition, so that those transitions will be most likely that conform most closely to the Principle. Therefore, *if* in the iodine molecule the curves are related as in set I of Fig. 1, a molecule initially not vibrating will most readily absorb light that carries it into states of high vibration, or to states higher than the energy needed for dissociation, resulting in photochemical dissociation of the molecule.

It was, of course, not a very difficult step to recognize that if the molecule is vibrating initially, then the Principle asserts that transitions are favored to those states which require least instantaneous adjustment of the relative position *and* momentum. Moreover, it was intuitively felt that the electronic transition was sufficiently independent of the nuclear vibration that it was equally likely to occur at any phase of the nuclear vibratory motion. This is not self-evident and perhaps not exactly true. It might be, for example, that the electron jump is stimulated by vibratory motion in the molecule, perhaps in such a way that it is more likely to occur at a phase of maximum relative velocity. However, this is not the case. With all times of electron jump equally likely, the most favored vibrational transitions will be those associated with the turning points of the nuclear vibration; for, as the nuclei move slowly here, a larger fraction of the time is spent in such regions.

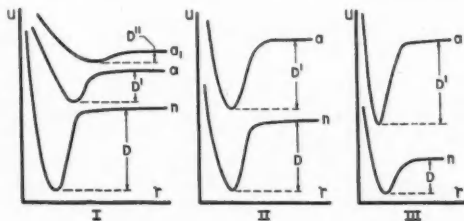


FIG. 1. Potential-energy curves from Franck's Faraday Society paper (reference 2).

² J. Franck, *Trans. Faraday Soc.* 21, 536 (1925). The paper in which the Principle was first advanced to explain the photochemical dissociation of iodine vapor.

³ E. U. Condon, *Physical Rev.* 28, 1182 (1926). First application of the Principle to intensity distribution in band systems.

This idea led to Fig. 2, taken from my 1926 paper,³ which indicates that there are two most favored vibrational quantum-number changes associated with nonvanishing values of the initial vibrational quantum number.

The proof of the pudding lay in the fact that a number of band systems were well analyzed, and therefore it was possible to put down the potential-energy curves in their approximately correct form and relative location. Near the minimum a curve is parabolic:

$$V(r) = \frac{1}{2}k(r-r_0)^2 + \dots$$

One can get the value of k from the vibration frequency ν , since, if μ is the reduced mass, $2\pi\nu = (k/\mu)^{1/2}$; and one gets r_0 from the moment of inertia, which in turn is given quantitatively from the observed rotational energy levels. Evidently in Fig. 2, when there is little change in ν or r_0 , the curves are similar and lie directly over each other, so the Principle requires small or zero change in the vibrational quantum number. But if there is a big change in r_0 , the curves are widely displaced and the Principle requires large changes in the vibrational quantum number. Both cases were found in the data available in 1926. The Principle was triumphant in that those with large changes in vibrational quantum number were correctly correlated with large changes in the equilibrium internuclear distance, and *vice versa*.

It must be remembered that the intensity data available were simply rough estimates of plate blackening, uncorrected for variations in plate sensitivity over the rather great range of wavelength involved in some band systems. Nevertheless, there was no doubt of the essential correctness of the Principle.

Curiously enough, my calculations indicated a bad disagreement with the facts for iodine, the very molecule which led Franck to his qualitative discovery of the idea. This caused me much worry, and I searched for an error in calculation a long time before sending in the 1926 paper with such a bad discrepancy for iodine. About a year later the error was discovered by Professor Wheeler Loomis, then of New York University. I had used the value of the moment of inertia for the 26th vibrational state in the electronically excited level, which is involved in the fluorescence

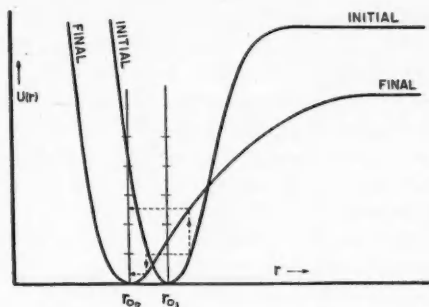


FIG. 2. Graphical construction of favored transitions from my 1926 paper (reference 3).

spectrum of iodine, thinking it was the moment of inertia of the nonvibrating molecules. When this error was corrected, iodine agreed with the Principle as well as the others.⁴

The faculty of the University of California was broad-minded enough to accept the paper as a doctor's thesis. However, it is interesting to note that orthodox theoretical physics was so tied to Bohr's correspondence principle at the time that the referee for the *Physical Review* was reluctant to recommend the paper for publication. He felt it could not be correct because it did not go at the problem in terms of Fourier amplitudes of the classical motion.

But the whole story was not given by the simple classical picture of 1926. The main puzzle might be stated thus: How exact is the Principle? or, what determines the extent of its inexactness?

The potential-energy curves lie in a definite location, so the Principle leads to a rather definite indication of most favored transition. Actually, although the predictions of the most favored transition agreed well with the facts, there was no indication of why other transitions could occur or how to calculate their intensities. This was a serious shortcoming of the theory, which Professor Birge did not hesitate to point out to me at the time.

A year later, in 1927, we all became familiar with Heisenberg's uncertainty principle as one of the broad implications of quantum mechanics. But in 1926 it did not occur to anyone that there was anything basically wrong with talking about

⁴ E. U. Condon, *Physical Rev.* **32**, 858 (1928). Development of quantum-mechanical formulation of the principle, including basic idea of internal diffraction.

definite values of both nuclear position and momentum being carried over without alteration in an electronic transition. Looking back on it later one can see that it might have been thought of, because statistical mechanics was dealing in terms of elementary cells of finite extension in phase space and the uncertainty principle is closely related to that. However, as elsewhere, hindsight is often better than foresight in theoretical physics.

Quantum-Mechanical Formulation

All that I just described happened in the spring of 1926. Matrix mechanics had been developed, but the matrix calculus was so difficult mathematically that it was hard to put any physical insight into it.* Schrödinger's famous series of papers on wave mechanics was just appearing; but it was not until the fall of 1926 that Born first recognized the probability interpretation of $|\psi|^2$ which collision problems forces on us, as contrasted with the more hydrodynamic views which are possible alternatives for closed systems.

I had the good fortune to be sent by the International Education Board from Berkeley to Göttingen in the fall of 1926 and thus was able to plunge into the study of the new quantum mechanics at one of the few points of high concentration of original thinking in this field. Great ideas were coming out so fast that period (1926-27) that one got an altogether wrong impression of the normal rate of progress in theoretical physics. One had intellectual indigestion most of the time that year, and it was most discouraging.

Besides studying the current papers I tried to solve the basic problem in quantum mechanics that underlies all molecular dynamics—the basic justification for the method of first working out the electronic states for fixed positions of the nuclei, so that the electronic energy levels depend parametrically on the nuclear coordinates and serve as the potential-energy functions for determination of the nuclear motions.

The justification is, of course, connected with

* In fact, I remember very well, in the fall of 1926 at Göttingen, that Professor David Hilbert said to his class in this connection, "die Physik wird zu schwer für die Physiker!"

the smallness of electron mass relative to nuclear mass, but I never could see how to work it out. Later the problem was handled in a basic paper by Born and Oppenheimer⁵ which, however, I have never felt that I properly understood. But Born is such a great master at seeing all physics in terms of "Entwicklungen nach einem kleinen Parameter Kappa" that I suppose it must be all right. The paper of Born and Oppenheimer is among those difficult ones that are more often cited than read.

These studies did, however, serve to make clear the place of the Principle in relation to the general ideas of quantum mechanics. Although this topic was treated in a short paper written from Göttingen, it was not properly handled until the fall of 1928 in a paper⁴ written in Princeton. The essence of the argument is that the wave function for a diatomic molecule is approximately of the form of a product of an electronic wave function, in which the nuclear coordinates appear as parameters, and of a wave function for the nuclear motion, say

$$\Psi = u(x_e, x_n) \cdot v(x_n),$$

where x_e stands symbolically for the electron coordinates and x_n for the nuclear coordinates. Therefore, the matrix element for a quantity like the dipole moment, which determines the radiation transition probabilities, is of the form

$$M_{12} = \int_e \int_n \Psi_1 M(x_e, x_n) \Psi_2 dx_e dx_n \\ = \int_n \bar{v}_1(x_n) M_{12}(x_n) v_2(x_n) dx_n,$$

where

$$M_{12}(x_n) = \int_e \bar{u}_1(x_e, x_n) M(x_e, x_n) u_2(x_e, x_n) dx_e.$$

The quantity $M_{12}(x_n)$ is the matrix element of the dipole moment calculated by regarding the nuclear coordinates as parameters of the electronic problem rather than as dynamical coordinates. It is characteristic of the electronic jump in question and so is the same for all the vibrational transitions of a band system.

⁵ M. Born and J. R. Oppenheimer, *Ann. Physik* **84**, 457 (1927). Basic quantum-mechanical justification of use of electronic potential-energy curves to determine nuclear motion in molecules.

The nuclear wave functions $v(x_n)$ may be quite complicated for a polyatomic molecule, but for a diatomic molecule they are of the form of a function of the radial coordinate $R(r)$ multiplied by a spherical harmonic for the simple motion under conservation of angular momentum.

It is hard to say much in general about $M_{12}(x_n)$. In fact, no explicit calculation of an example has been made even yet. But it is natural to suppose that it will be a slowly varying or smooth function of r over the small range of r in which the radial wave functions have appreciable value.

The radial functions $R(r)$ are rather closely related to the corresponding classical vibratory motion, according to the general correlation provided by the Wentzel-Brillouin-Kramers approximation. According to this, the wave function is of the form

$$R(r) = \frac{1}{4(p)^{1/4}} \cos \left[(2\pi/h) \int^r p dr + \alpha \right]$$

within the range of the classical motion, and falls off rapidly to zero outside the range of the classical motion, being dominated by a factor of the form

$$\exp \pm \left[(2\pi/h) \int^r |p| dr \right],$$

where p is given by

$$(1/2\mu)p^2 + V(r) = W,$$

where $V(r)$ is the effective potential energy of the radial motion, including the effects of the rotational energy;

$$V(r) = V_0(r) + \frac{\hbar^2 J(J+1)}{2\mu r^2},$$

when J is the rotational quantum number and $V_0(r)$ is the potential-energy function applicable in the case of a nonrotating molecule.

It is easy to see qualitatively that the value of an integral of the form

$$\int R_1(r) R_2(r) dr,$$

with wave functions given by the Wentzel-Brillouin-Kramers approximation, is in the main determined by the conditions set up in the

classical formulation of the Principle as given in my 1926 paper:

(1) The integral will be small unless the wave functions overlap, that is, unless there is little sudden change in the internuclear distance called for in the transition;

(2) The integral will be small if the wave functions are related in such a way that a nonoscillatory part of one wave function overlaps a rapidly oscillatory part of the other, for this means that the electron jump would have to be accompanied by a considerable change in the radial momentum.

This formulation goes considerably further than that of the 1926 paper in three main respects:

(i) It follows in a definitely deductive way from a well-founded theory whose other successes have given it a definite standing as an adequate basis for atomic mechanics;

(ii) It gives, in principle, definite results for the relative strength of each possible quantum transition; it therefore goes beyond the simple classical picture based on the potential-energy curves and provides a basis for calculating the extent to which the less favored transitions actually do occur;

(iii) It leads to the prediction of some specific effects arising from the wave nature of matter for which I want to introduce the term "internal diffraction;" these will be dealt with at length in the next section.

When it comes to setting up explicit formulas for the integrals $\int R_1(r) R_2(r) dr$ it is not possible to get results of very general applicability. The natural thing to do is to start with the approximation that the initial and final potential-energy functions are Hooke's law parabolas, differing as to force constant and as to equilibrium separation. The formulas on this supposition were set up and evaluated by Hutchisson.⁶ However, the results are rather complicated.

But a more severe limitation is that the cases which are most interesting physically are those in which there is a fairly large change in the equilibrium separation occasioned by the electronic transition, so that large changes in the vibrational quantum number occur. In such cases the

⁶ E. Hutchisson, *Physical Rev.* **36**, 410 (1930); **37**, 45 (1931). Explicit calculation of integrals for transitions using harmonic-oscillator wave functions.

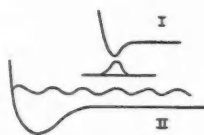


FIG. 3. Wave functions as related for internal diffraction in continuous spectra, from my 1928 paper (reference 4).

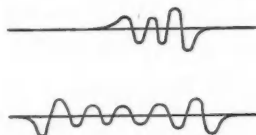


FIG. 4. Rough sketch of wave functions showing sensitivity of integral to relative location of nodes.

wave functions have to be known with fair accuracy at some distance away from the equilibrium separation. The actual force law is not parabolic, so that the harmonic-oscillator wave functions are no longer good approximations over an important part of the range of the coordinates.

Since the actual departure from the harmonic law is different for each molecule, general formulas are of small value. The only proper test of the theory is therefore explicit calculation based on the specific facts in particular cases.

Internal Diffraction

I particularly want to stress the specifically nonclassical, or wave-mechanical, features of the theory, for it seems to me they should be more widely known. These features seem to me to be more than mere quantum-mechanical refinements of detail, as they have sometimes been presented in books. On the contrary, internal diffraction is to me just as real and forceful a proof of the wave nature of nuclear motion as any of the basic external diffraction experiments, such as those in which a beam of electrons or neutral hydrogen atoms is diffractively scattered by a crystal.

Figure 3, from my 1928 paper,⁴ presents one particular situation in which internal diffraction might rise. It does not matter whether curve *I* is above or below curve *II*: if *I* is above, the phenomenon will appear in emission, while if *I* is below, then it will appear in absorption. The wave function of the lowest vibrational level in state *I* will be approximately a Gauss error function as shown. A radial wave function for a

typical energy value in the continuum above the dissociation limit of *II* will look like the one sketched in the figure. As the energy of the final state is increased the radial wave function will vary in this way: its first loop will have a fairly constant relation to the turning point for the classical motion; but, as the de Broglie wavelength slowly decreases, the nodes become more closely spaced, so that the phase of the quasistationary wave located under the center of the Gauss error wave function gradually changes. Clearly, when a loop of the sinoidal wave function is under the center of the Gaussian wave function we shall get a large value of the integral which governs radiative transition. But when a node is under the Gaussian wave function, the integral is very small.

In this way the initial-state wave function is able to "see" the wave nature of the final-state wave function. Transitions can occur to the energies of the continuum for which a loop is under the initial wave function and are much weaker when they occur to the energies having a node there. The result is a rippling variation in the intensity of the continuous spectrum. Such a manifestation of the wave detail of the ψ function I call *internal diffraction*, in analogy with external diffraction, which is also determined by a phase relation between initial and final wave functions.

The term internal diffraction is a convenient one to use more generally to describe specific

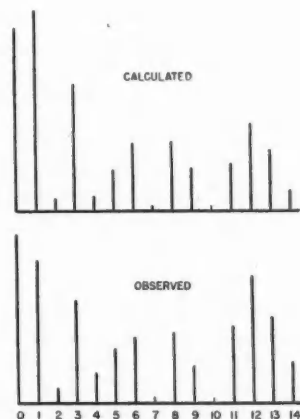


FIG. 5. Brown's calculations on sodium fluorescence bands (reference 7).

variations of the transition probabilities away from those indicated by the Principle after it is fuzzed out by uncertainty principle requirements. In the more general case we may suppose that two wave functions for initial and final states are related approximately as in Fig. 4. Evidently, the exact value of the integral of the product of two such functions is quite sensitive to the exact relative positions of the nodes and loops. The integrand is itself a roughly oscillatory function, being positive where the factors are of the same sign and negative where they are of opposite signs. Hence the value of the integral is quite sensitive to the average relative phase of the oscillations in the wave functions of initial and final states. This is a specifically quantum-mechanical effect, and definite results calculated from it afford a sensitive test of the reality of de Broglie waves associated with the nuclear motions in the molecules.

Although it did not seem feasible to make sufficiently accurate calculations to check the point at the time, this view was put forward in 1928 to explain the intensity fluctuations observed by Wood in the fluorescence spectrum of iodine vapor. On excitation by the green line of mercury, diatomic iodine molecules are raised to the 26th vibrational level of an excited electronic state from which they emit a long series of fluorescence

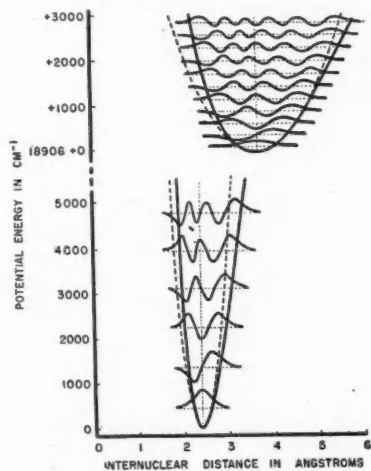


FIG. 6. Potential curves and wave functions used by Gaydon and Pearse in calculations on rubidium hydride (reference 8).

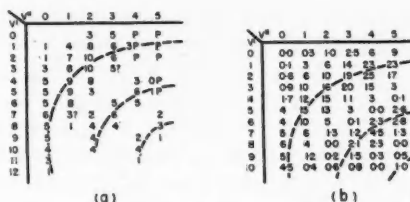


FIG. 7. Intensities for rubidium hydride as given by Gaydon and Pearse: (a) observed visual estimates on a scale of 10; (b) calculated, reduced to a scale of 25.

doublents on jumping to various vibrational levels of the normal electronic state. These doublets form a series the intensity of which varies in quite an irregular way, namely: 10, 9, 1, 9, 3, 8, 8, 2, 9, 0, 8, 3, 2, 7, 0, 7, 0, 2, 1, 0, ... All of these transitions are permitted by the approximate form of the Principle. Undoubtedly the fluctuations are due to the particular phase relations of the nodes in the initial and final states, although because of the high quantum numbers involved an explicit calculation to test this statement would be very laborious and has not been made.

But the lack has been pretty well supplied by two explicit calculations which will now be mentioned. A similar irregular variation of intensity occurs in the fluorescence bands of Na_2 . W. G. Brown⁷ in 1933 calculated the approximate integrals and obtained the comparison of observed and calculated intensities that is shown in Fig. 5.

In 1939 similar calculations were made by Gaydon and Pearse⁸ the band system of rubidium hydride. Figure 6, from their paper, shows the potential-energy functions and wave functions that they used. The dashed parabolas show the Hooke law approximation to the potential energy, and the full curves show the more accurate potential-energy functions inferred from the levels. The wave functions used were obtained by a reasonable approximate transformation from harmonic-oscillator wave functions rather than by numerical integration of the wave equation.

⁷ W. G. Brown, *Zeits. f. Physik* **82**, 768 (1933). Explicit calculation of internal diffraction intensity variations in fluorescence bands of diatomic sodium molecule.

⁸ A. G. Gaydon and R. W. B. Pearse, *Proc. Roy. Soc. A* **173**, 37 (1939). Careful calculation of transition probabilities in band spectrum of rubidium hydride, showing internal diffraction effects quantitatively.

The result of their calculations is shown in Fig. 7. The table on the left gives observed values (estimates) and that on the right the results of calculation. The main parabolic locus of strong bands is the familiar result of the classical Principle. The secondary weaker loci are the result of special phase relations between initial and final wave functions, that is, internal diffraction.

Molecular Hydrogen

Diatomic hydrogen is a particularly interesting molecule since it has the simplest structure, one that is simple enough to permit of making some fairly accurate theoretical calculations. These calculations, initially made by Heitler and London,⁹ provided a host of new viewpoints applicable in general to chemistry. In particular, they cleared up the basic nature of the electron-pair or homopolar valence bond. There has to be some limit to the scope of this review, so my remarks will be confined to some interesting points that grew out of the application of the Principle to the theoretical potential-energy curves of Heitler and London, and of Burrau¹⁰ for H_2^+ .

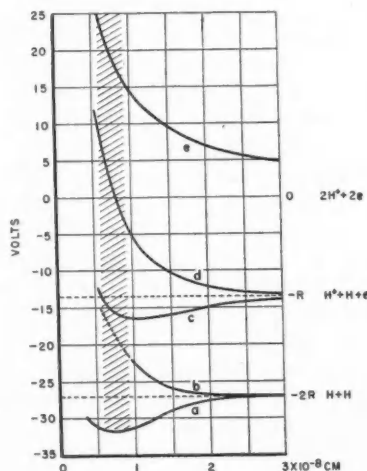


FIG. 8. Potential-energy curves for molecular hydrogen, from Bleakney's 1930 paper (reference 13).

⁹ W. Heitler and F. London, *Zeits. f. Physik* **44**, 455 (1927). First quantum-mechanical calculation of potential-energy curves of molecular hydrogen.

¹⁰ O. Burrau, *Kgl. Danske Videnskab. Selskab. Math-fysiske Medd.* **7**, 14 (1927). First calculation of potential-energy curves for normal state of ionized hydrogen molecule.

The application of the Principle to molecular hydrogen was first made by Winans and Stueckelberg¹¹ in 1928. They recognized that radiative transitions from the excited triplet levels on to the repulsive-force triplet sigma ($^3\Sigma$) state of Heitler and London could account for the very extensive ultraviolet continuous spectrum of molecular hydrogen. At about the same time a paper written with Smyth¹² gave some interpretations of observed critical potentials in terms of the repulsive curves. These two papers may be said to have provided the first evidence of the physical reality of the repulsive-force curves given by quantum mechanics.

A much more convincing proof of the physical reality of such electronic states in molecular hydrogen was given in 1930 by Bleakney¹³ when we were together at the University of Minnesota.

In the older data on critical potentials in hydrogen a value close to 31.5 v had usually been reported. As this agrees quite well with the minimum value of the energy needed to strip the two electrons off a hydrogen molecule, it was assumed to correspond to the process of total dissociation of the hydrogen molecule into two protons and two electrons.

But how could this be? According to the Principle, the colliding electron would have to strip off the two electrons, leaving the two protons at essentially the same distance apart as they are in the normal molecule. At this distance, the two protons have a Coulomb interaction energy of about 20 ev; therefore the process of total dissociation would require about 51.5 v instead of 31.5 v. Subsequently, of course, the two protons would fly apart under their mutual repulsion, each of them getting 10 ev of energy.

The situation is shown in Fig. 8, taken from Bleakney's 1930 paper.¹³ Curves (a) and (b) are derived from the basic Heitler and London solution for the mutual interaction of two normal hydrogen atoms. Curve (c) is from Burrau's

¹¹ J. G. Winans and E. C. G. Stueckelberg, *Proc. Nat. Acad. Sci.* **14**, 867 (1928). Interpretation of the continuous ultraviolet emission spectrum of molecular hydrogen.

¹² E. U. Condon and H. D. Smyth, *Proc. Nat. Acad. Sci.* **14**, 871 (1928). Interpretation of the critical potentials of molecular hydrogen.

¹³ W. Bleakney, *Physical Rev.* **35**, 1180 (1930). Experimental proof that some ions formed on electronic impact with molecular hydrogen have kinetic energy.

calculation for the normal state of H_2^+ . Curve (d) is the analogous repulsive-force state in H_2^+ as calculated by Morse and Stueckelberg.¹⁴ Finally, curve (e) is simply e^2/r , the potential-energy curve of the molecule H_2^{++} .

As soon as these curves were drawn some interesting conclusions were evident. In the first place it is clear that the Principle does not allow us to dissociate the molecule simply by striking it with an electron of the minimum necessary energy, in this case about 4.4 v, for that would require a large change in position or momentum of the nuclei to accompany the electronic transition.

However, at about 11 v it ought to be possible to dissociate the molecule by electron impact by inducing transitions from state (a) to state (b). This transition is forbidden for light absorption but allowed for electron impact. The excited H_2 molecules would immediately fly apart, giving two normal hydrogen atoms each with about 3.5 ev of kinetic energy. As both products are neutral they could not be observed in a mass spectrometer. Observations have been made that indicate a rapid clean-up—that is, adsorption on the walls—of hydrogen, which could be due to formation of atomic hydrogen at about this energy, but not at lower voltages.

Considering state (c), the change in equilibrium separation indicates that the transition from H_2 to normal H_2^+ not vibrating is quite unlikely, and that it is likely that there will be a good yield of $H+H^+$ by direct transition from the normal state to the part of (c) lying above the dissociation limit. This was observed to be the case.

Coming now to curve (d), we notice the explanation of the 31.5-v critical potential. It has nothing to do with total dissociation; that would require transitions to curve (e). The 31.5-v potential is in reality the transition, indicated by the Principle, to curve (d). If this is the correct explanation, then the H^+ ions formed in this process should have about 6.5 ev of kinetic energy. By an appropriate use of retarding fields in his mass spectrometer, Bleakney could show that this was so. It constituted the first instance of observation of molecular-ion fragments formed with kinetic energy and the most direct and un-

ambiguous proof yet found of the physical reality of the repulsive-force molecular states predicted by quantum mechanics.

Later more careful data were taken by Tate and Lozier,¹⁵ establishing that molecular-ion fragments can be formed with kinetic energy in other molecules as well, although here the quantum-mechanical calculations are too difficult to permit detailed predictions to be made.

There is another point in this connection, which was studied briefly by Hipple in 1936, that provides another interesting example of a quantum-mechanical phase of the Principle. Consider again the transitions from curve (a) to curve (c), induced by electron impact. The minimum of curve (a) is so related to (c) that the most favored transitions are those leading to H_2^+ molecules in a rather high vibrational state. If we bombard H_2 with, say, 18-v electrons, we will get some H^+ but transitions to H_2^+ are more favored.

What happens if we use deuterium instead of hydrogen? Theory tells us that the potential-energy curves are quite accurately the same in the two isotopic molecules. Theory also tells us that the wave function of the normal zero-vibrational state in D_2 will be narrower than in H_2 ; because of its greater mass it behaves more classically. Hence we can predict a lower yield of D^+ relative to D_2^+ at 18 v in D_2 than of H^+ relative to H_2^+ in H_2 . Hipple tried this and found a considerable effect. It would be interesting to get accurate data on this and to attempt a precise calculation.

In 1931 Finkelberg and Weizel tried to get definite information on the shape of the basic Heitler-London repulsive-force curve by an interpretation of data on the potentials needed to excite different parts of the continuous spectrum of molecular hydrogen. Such a procedure involved using the Principle in a more precise way than was ever intended, as was pointed out in 1936 by Coolidge, James and Present.¹⁶ They did a very careful job, first on an improved calculation from theory of the repulsive-force curve, then on numerical integrations of the radial wave-function products occurring in the theory,

¹⁵ J. T. Tate and W. Lozier, *Physical Rev.* **39**, 254 (1932).

¹⁶ A. S. Coolidge, H. M. James and R. D. Present, *J. Chem. Physics* **4**, 193 (1936). Careful discussion of application of the Principle to continuous spectrum of molecular hydrogen.

¹⁴ P. M. Morse and E. C. G. Stueckelberg, *Physical Rev.* **33**, 932 (1929). First calculation of repulsive-force potential-energy curve for ionized hydrogen molecule.

in order to get a proper treatment instead of the usual rough graphical construction from the potential-energy curves.

The work of Coolidge, James and Present brought out some fairly good evidence that in this case one cannot treat the electronic dipole-moment matrix component as constant. This meant that the rough assumption I had made earlier in order to get the main idea straightened out needed to be improved. Because of this point they were led to say in the abstract that "It is concluded that the Franck-Condon principle leads to results definitely incompatible with observations." I cannot let that pass, even ten years later, without a word of protest. They can say that I misused the Principle if they like, but not that the Principle is incompatible; for the Principle is properly to be judged by its correct quantum-mechanical formulation, which they so beautifully worked out, rather than by the rough criteria I gave as approximate rules.

The whole situation with regard to these special properties of molecular hydrogen seems to me to be highly satisfactory.

Conclusion

In the past decade, despite both the competitive attraction of nuclear physics and the wasteful interruptions of the war, many more band systems have been analyzed so that today there are a large number of instances of electronic transitions in a wide variety of molecules, all of which conform nicely to the semiquantitative

intensity relations given by the approximate form of the Principle. Nevertheless, the fact remains that the Principle has never been subjected to a really severe test in which intensities are carefully measured by good photographic photometry and compared with those calculated from highly accurate wave functions.

I will conclude this review by merely mentioning one other application. The Principle has also contributed a good deal to the elucidation of many points concerned with predissociation of molecules.¹⁷ Here, except for a few cases in which feebly bound molecules can dissociate by rotational instability aided by potential barrier leakage, we are concerned with radiationless transition between two electronic states of equal energy. The Principle acts restrictively here also, in that such transitions cannot occur if they would require too much readjustment of the nuclear motions.

The Principle also applies "in principle" to polyatomic molecules,¹⁸ although here the situation is much more complicated than in the atomic case. Not only are the band systems much more complicated, so that thus far little progress has been made in analyzing them, but also the potential-energy curves have to be replaced by potential-energy surfaces, which are not well known either empirically or theoretically.

¹⁷ L. A. Turner, *Zeits. f. Physik* **68**, 178 (1931). Application of the Principle to predissociation.

¹⁸ J. Franck, H. Sponer and E. Teller, *Zeits. f. physik. Chemie* **18**, 88 (1932). Application to predissociation in polyatomic molecules.

*We always think of fascism as of something solid, palpably different from our forms of life. It would be more correct to compare it with a gas which can be put into any container regardless of its shape. And once you get into the habit of living amidst a moderate amount of stink you won't notice it when you become completely poisoned. The danger is not that we may wake up one morning to find a fascist world; that would be easy to prevent. The danger is that we went to bed the previous night in a world which was already turning fascist without our noticing it. . . .—ARTHUR KEOSTLER, *The yogi and the commissar* (Macmillan).*

Sample Illustrations of Physical Principles Selected from Physiology and Medicine

L. A. STRAIT AND V. T. INMAN
University of California, San Francisco 22, California

AND

H. J. RALSTON
College of Physicians and Surgeons, San Francisco 3, California

THE teaching of the principles of physics requires, beyond the statement of principles, the use of illustrative examples. In great part such examples are chosen, for obvious reasons, from experience in the engineering fields. If, in addition, numerous examples from the biological and medical fields were to be included it seems reasonable that the desire of the premedical student to learn physics would be stimulated. The Association's Committee on the Teaching of Physics to Biological and Premedical Students¹ has adopted as one part of its program the preparation of appropriate examples and source material which might be used to supplement the conventional elementary course in physics.

From the practical point of view there are two major aspects to achieving this end: (i) the discovery of appropriate illustrative material; (ii) the reduction of the material to usefulness in physics instruction. The latter requires that the emphasis remain on the teaching of physics and that the biological scientific knowledge assumed be the minimum necessary for the visualization of the physical principles. That type of illustration which requires the least biological prefacing is then the most desirable.

Pressure, Force, Work, Power and Efficiency Illustrated by Heart Action

The heart maintains the life process of circulating the blood by cyclic muscular contractions and relaxations. The action of the heart essential to our purpose may be briefly presented in a single picture requiring but minor explanation. Figure 1² shows clearly the "force-pump" action of the heart due to a contraction of the two

chambers (ventricles) after they are filled with blood in the previous relaxed part of the heart cycle. The contraction of the right ventricle pumps blood to the lungs only; the left ventricle pumps blood throughout the entire body. The contraction part of the cycle for the left ventricle may be reduced further to that of the force stroke of a piston in a cylindrical chamber, as in Fig. 2.

With this degree of physiological introduction one may use the action of the heart as a biological source of the following examples of elementary physical principles.

(1) *Force and pressure: maximum force exerted upon the left ventricle by the blood.* Referring to Fig. 1, if the inner surface of the left ventricle of the heart is taken to be 82 cm² and if the contained blood acquires a maximum pressure of 120 mm of mercury because of the contraction of the chamber, evaluate the total force in dynes and in pounds exerted upon the inner surface of the ventricle at the instant of maximum pressure. *Ans.* 29.4 lb.

(2) *Work ($W = P\Delta V$): work of the left ventricle of the heart.* Referring to Fig. 2, assume that the left ventricle of the heart is analogous to a single piston and chamber, and that 70 cm³ of blood³ are ejected against an average intraventricular pressure of 105 mm of mercury. Calculate the

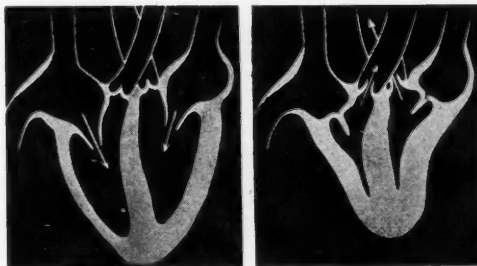


FIG. 1. Relaxation and contraction of the chambers of the heart. The arrows indicate the direction of blood flow. [Courtesy of Encyclopaedia Britannica Films, Inc.]

¹ L. L. Barnes, L. I. Bochstahler and L. A. Strait, *Am. J. Physics* 14, 338 (1946).

² A. Carlson and V. Johnson, *The machinery of the body* (Univ. of Chicago Press, 1937), p. 132, Fig. 40. This figure is taken from the Encyclopaedia Britannica film, "The heart and circulation."

³ W. Howell (Ed. by Fulton), *Textbook of physiology* (Saunders, 1946), pp. 795-799; E. H. Starling, *Principles of physiology* (Lea and Febiger, 1936), pp. 714-716.

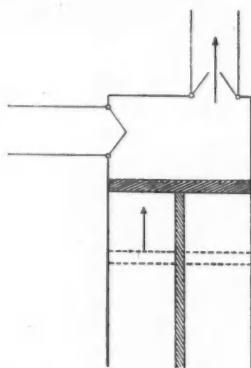


FIG. 2. Schematic equivalent of the contraction portion (the systole) of the heart cycle for the left ventricle (chamber) of the heart.

work done by the ventricle in a single contraction. *Ans.* 9.8×10^6 ergs.

(3) *Power: power developed by the left ventricle of the heart.* Assuming one beat of the heart per second, calculate the useful power which the left ventricle develops (in the contraction portion of the cycle) in watts and in horse power. *Ans.* 0.98 w; 1.31×10^{-3} hp.

(4) *Work: estimate of minimum work done by the left ventricle in a life span of 70 yr.* Assuming that a minimum of 0.001 hp is expended by the left ventricle in doing the work of circulating the blood, calculate the total work

done in 70 yr. This assumes a man at rest in bed for 70 yr, an idealization that students would appreciate. *Ans.* 1.2×10^9 ft lb.

(5) *Efficiency: the heart as an engine.* The heart may be considered to be an engine using oxidative chemical reactions as a source of energy. The amount of energy supplied to a body tissue by the carbohydrate, fat and protein fuel is related to the amount of oxygen used in the reactions. It has been determined by experiment that each liter of oxygen used by body tissues corresponds to a liberation of approximately 4800 cal of energy.⁴

If a heart isolated from a dog were observed to consume 14.2 cm^3 of oxygen per minute,⁵ calculate the efficiency of the heart, assuming the power expended by the heart in useful work to be $120 \times 10^6 \text{ erg/min}$. By efficiency here is meant the ratio of work done by the "piston" and the energy supplied by the fuel. *Ans.* 4.2 percent. The efficiency of the heart in the intact animal may be as high as 25 percent.

Although these examples require considerable simplification and idealization of the complex biological systems from which they are taken, they do not seem to do violence to the biological facts (if the simplifying assumptions are stressed) to a greater extent than do engineering examples that involve weightless members, frictionless surfaces, and so forth.

As in other kinds of illustrations there remains a choice as to the degree of simplification desired. A balanced bibliography should accompany each example, so that those interested beyond the use of the "reduced" illustration for purely pedagogic purposes may satisfy their curiosity or seek to extend the illustration or examples. In the case of the heart, a somewhat more realistic and advanced treatment would recognize that the flow is that of a viscous fluid and that the pressure is not constant during the contraction and is generally measured in the aorta (rather than in the ventricle) as a lateral pressure. The same example may be readily extended to include an illustration of Bernoulli's equation for the energy of flow.

Variations of presentation can no doubt heighten the student's interest. For example, the billion foot pounds of work done by the "resting" heart in Example 4 might be compared with the work required to lift a half-dozen rail-

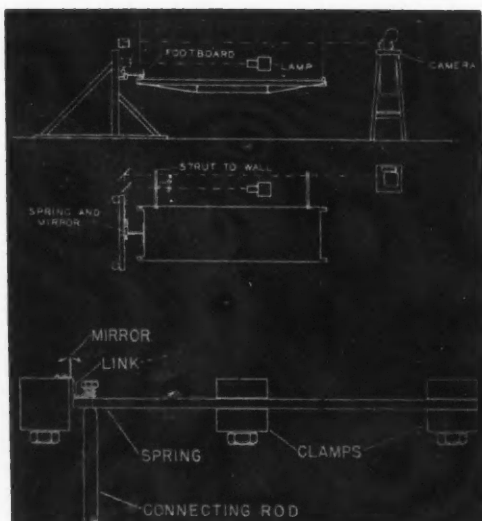


FIG. 3. Details of a ballistocardiograph used by Starr, *et al.* (reference 8). Side and top views of the suspended table are shown in the upper portion of the figure. The restraining spring and the optical lever are shown below.

⁴ W. Howell, reference 3, pp. 1091-1092; E. H. Starling, reference 3, pp. 454-455; O. Glasser, Ed., *Medical physics* (Year Book Publishers, 1944), pp. 215, 578.

⁵ Fahr, George and Buehler, *Am. Heart J.* 25, 211-241 (1943).

road cars fully loaded with coal to the height of the Empire State Building. The efficiency of the heart may be compared with that of a Diesel-type engine (30-35 percent), a gasoline engine (25-30 percent), and a steam locomotive (10 percent).

Beat of the Heart and Conservation of Momentum

If one stands motionless on the platform of a sufficiently sensitive balance he can observe minor pulsating fluctuations of the needle from its position of equilibrium. The frequency of the fluctuations will be observed to be correlated with the beat of the pulse. (With care this effect can be observed on the conventional bathroom "scales.") These motions reflect the effect of the action-reaction forces existing between the ejected blood and the rest of the body, and may be cited as an illustration of the principle of conservation of momentum, analogous to the recoil in the firing of a projectile. Because the blood upon ejection from the heart does not escape from the body, but in the process of circulation suffers many changes of momentum, the body will experience a series of slight movements as a consequence. (The over-simplified analogy to the recoil of a gun and its projectile might be made to resemble this case more closely if the gun were mounted on a small boat and were firing a shot that ricocheted in a complex manner among targets mounted on the boat.)

Sensitive methods of recording and of measuring these slight body movements have been devised⁶⁻⁸ and used for physiological study of the action of the heart. Such devices are called *ballistocardiographs*. The ideal ballistocardiograph would be a device that would record the reaction motions of the body if it were suspended freely in space. Suspended chairs and tables (swinging freely or constrained to head-footward motions only), a rigid table and seismograph, and other forms of apparatus have been used (Fig. 3). A novel form of apparatus was the ingenious use, on an expedition to Pikes Peak,⁹ of a plank sup-

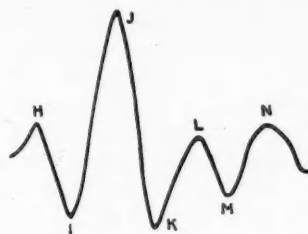


FIG. 4. Typical record of one heart cycle obtained with the ballistocardiograph. A downward deflection represents a footward movement of the table.

ported on piles of cork. The small displacements have been amplified optically and electronically for recording. A typical diagram of the most common form of normal record obtained by a ballistocardiograph for one heart cycle is given in Fig. 4.⁸

The detailed mechanical analysis of ballistocardiographs and the form of the records is complex. The effect of the mass of the supports and the constraining forces imposed on practical devices presents a rather advanced problem in analysis. This difficult problem has received

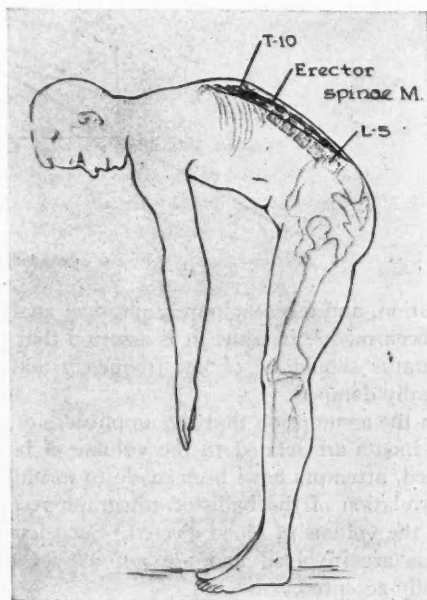


FIG. 5. Diagrammatic representation of the muscles used in raising the trunk. The tenth thoracic and fifth lumbar vertebrae are indicated.

⁶ W. Howell, reference 3, p. 298.

⁷ O. Glasser, reference 4, p. 575.

⁸ Starr, Rawson, Schroeder and Joseph, *Am. J. Physiol.* 127, 1-27 (1939).

⁹ Douglas, et al., *Phil. Trans. Roy. Soc. B* 203, 185 (1913).

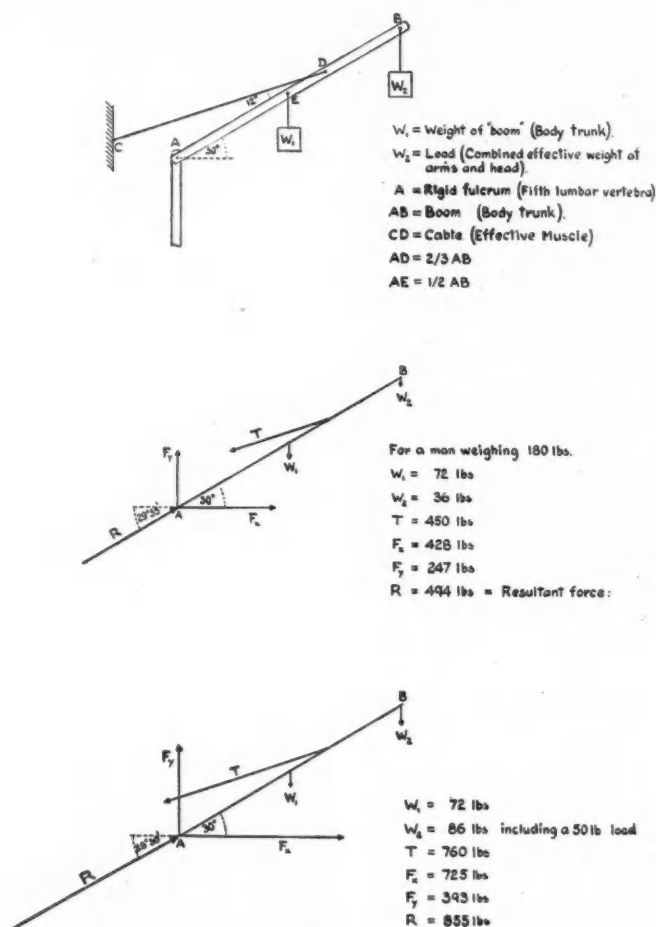


FIG. 6. Idealized mechanical model of Fig. 5.

attention, and a recent purely physical analysis has been made¹⁰ in which it is asserted that the apparatus should be of low frequency and be critically damped.

On the assumption that the amplitudes of the movements are related to the volume of blood ejected, attempts have been made to formulate a correlation of the ballistocardiograph records with the volume of blood ejected.⁸ Such formulations are involved and are not as yet universally accepted.

In the example of the ballistocardiograph it is apparent that there is ample latitude in the

degrees of complexity of presentation to stimulate the physics teacher as well as the student.

Muscular Mechanics and Equilibrium of Forces

The human body may be likened to a physical machine made up of many mechanical members; the joints are fulcrums, and the contracting muscles exert forces on the members analogously to simple cables. When the muscle forces balance the pull of gravity plus the applied load, the member is in equilibrium. The opportunity is thus provided for illustration of the physical principles of equilibrium with demonstration equipment which can be carried in and out of the

¹⁰ Nickerson and Curtis, *Am. J. Physiol.* **142**, 1 (1944).

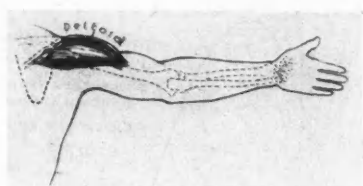


FIG. 7. Diagrammatic representation of the deltoid muscle, used in raising the arm.

classroom with no extra effort on the part of the instructor. (Furthermore, it is a subtle way of demonstrating that the instructor has human qualities—a matter of doubt to some students!)

That we are designed for speed of motion of the parts and conservation of bulk rather than for the incompatible maximum of strength or of mechanical advantage leads to some startling values of muscular force requirements, for example, for support of our limbs. The force on

the fifth lumbar vertebra which serves as a fulcrum for the spine in bending forward is an example. Figure 5 should make clear to the student the (elementary) anatomy involved, and Fig. 6 is an idealized illustration of a 180-lb man bending his trunk forward 60° from the vertical with arms hanging freely. Considering the fifth lumbar vertebra as a rigidly fixed fulcrum, how much force does it sustain? Assume that the head and freely hanging arms act as a single mass of weight, W_2 , and that their combined weight is 20 percent of body weight.¹¹ Then W_2 is 36 lb. Assume further that the weight of the trunk W_1 is 40 percent of the body weight, or 72 lb, and that the distance from the fifth lumbar vertebra to the point of action of W_2 is L . The "effective" erector spinalis muscle, when the body is so inclined, acts at an average angle of 12° with the spine.¹² The muscle may be considered to be attached at a point distant $2L/3$ (average posi-

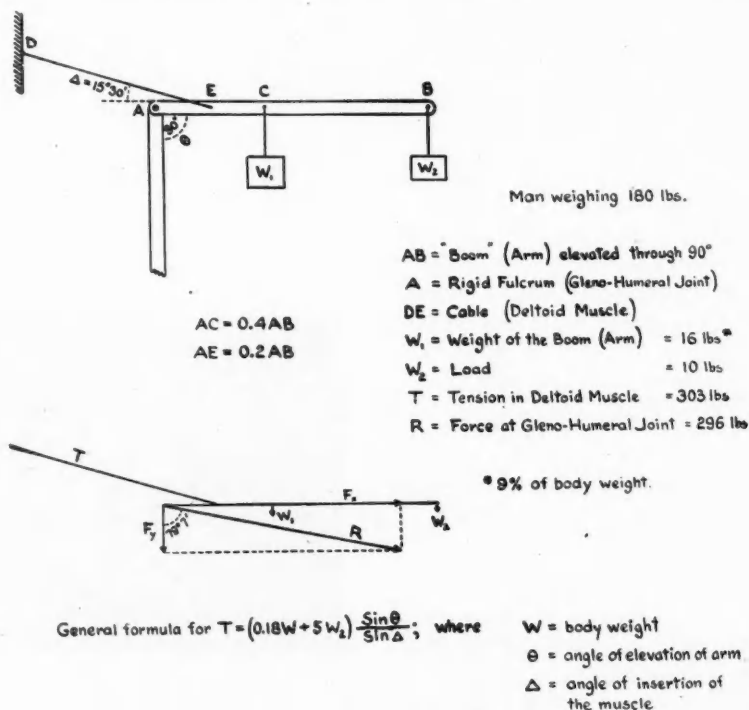


FIG. 8. Idealized mechanical model of Fig. 7.

¹¹ O. Fischer, *Wissenschaften* 25, 1-130 (1899).

¹² V. T. Inman, *Evaluation from x-ray analysis*, unpublished.

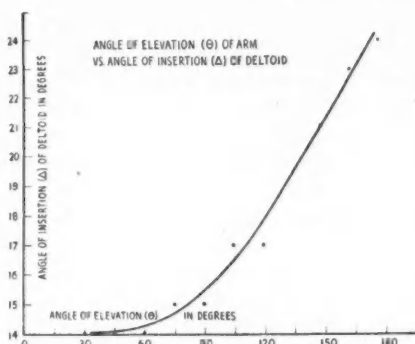


FIG. 9(a). The indicated angles have been determined from x-ray studies.

tion of insertion of the muscles) from the fifth lumbar vertebra.

The solution of this equilibrium problem shows that merely to compensate for gravitational forces on the unloaded trunk of a man weighing 180 lb, the tensile force in the effective erector spinalis muscle must be of the order of 450 lb, and a compressional force on the fifth lumbar vertebra of nearly 500 lb is required. When the man is loaded with 50 lb held in the hands, these forces become about 750 lb and 850 lb, respectively.

The student might be interested in trying to guess at some of the medical implications of the analysis of forces required. First calling attention to the magnitude of the force, it may be inferred that it is not too surprising that low back pains are the most common and that the fifth intervertebral disk is the one most frequently involved in disk rupture and early degenerative change. (The intervertebral disk acts like a shock absorber, not unlike a partially filled hot water bottle.) The structural weakness of design of man's lower back and the resulting back pains are, according to the anthropologists, the price man pays for having arisen from all fours to free his arms and hands. It may be noted that the resultant force is nearly parallel to the spinal column, making the strain on the disk primarily compressional with negligible shear. Injury is most likely to occur when shearing forces are introduced.

A proper appreciation of the mechanical factors involved is necessary for a full understanding of

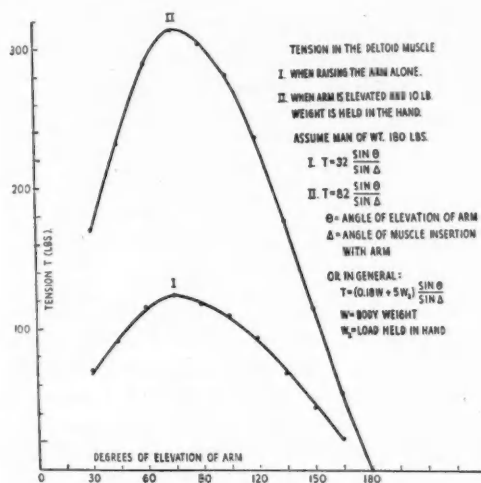


FIG. 9(b). Tensile force in the deltoid muscle.

the medical problem and may serve to encourage the student to recognize the usefulness of "theoretical" physical analysis as a tool.

Force in the deltoid muscle and reaction at the gleno-humeral joint arising from the elevation of the arm.—Another example of the surprisingly large force requirements of body muscles is presented in elevation of the arm. The pertinent data for a reasonably good approximation to the force are available, including the angle of insertion of the muscle (by x-ray analysis), and the center of gravity and relative weight of the arm.¹³ The

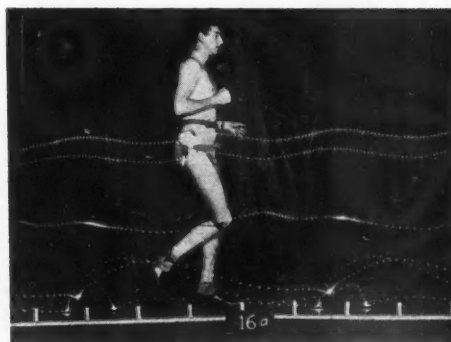


FIG. 10(a). Movements of various parts of the body during walking. Lights (see text) have been placed near the center of gravity of the body, at the upper end of the thigh (trochanter), and at the knee, ankle, heel and toe.

¹³ V. T. Inman, J. B. Saunders and L. C. Abbott, *J. Bone and Joint Surgery* 26, 1-30 (1944).

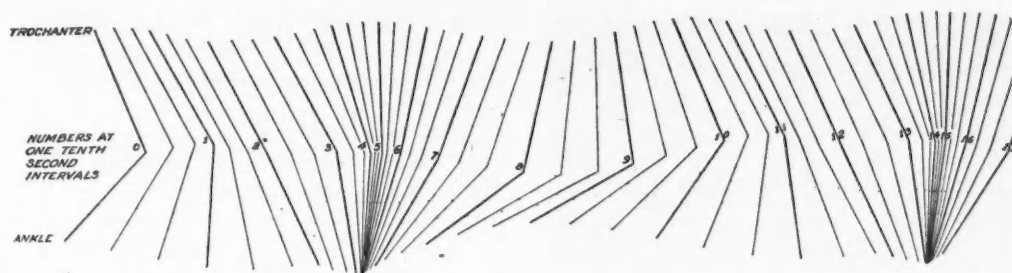


FIG. 10(b). Successive positions of the trochanter, knee and ankle, determined from Fig. 10(a).

anatomy involved is presented in Fig. 7, and the idealization and solution of the corresponding physics problem are given in Fig. 8, where two cases are considered. The first is with the arm unloaded, the second with a load of 10 lb in the hand. Figure 9(a) shows the experimentally determined function of the angle of insertion Δ , as a function of the angle of elevation θ . Figure 9(b) shows the calculated function of tensile force in the muscle while elevating the arm.

Kinematics.—The motions of our bodies in walking are more complex than the simple motions customarily studied in the elementary

physics courses. An ingenious method of studying these motions, used first by Otto Fischer,¹¹ not only is of interest as a simple experimental method of measuring space-time relationships of a motion, but also makes some aspects of the motion of the body and limbs readily understandable to the student.

Small lights are attached to the subject [Fig. 10(a)], who walks in darkness over a straight path. Photographs of the moving lights at successive, equal, small intervals of time are obtained by rotating a sector slit in front of the camera (30 rev/sec in the example given here).

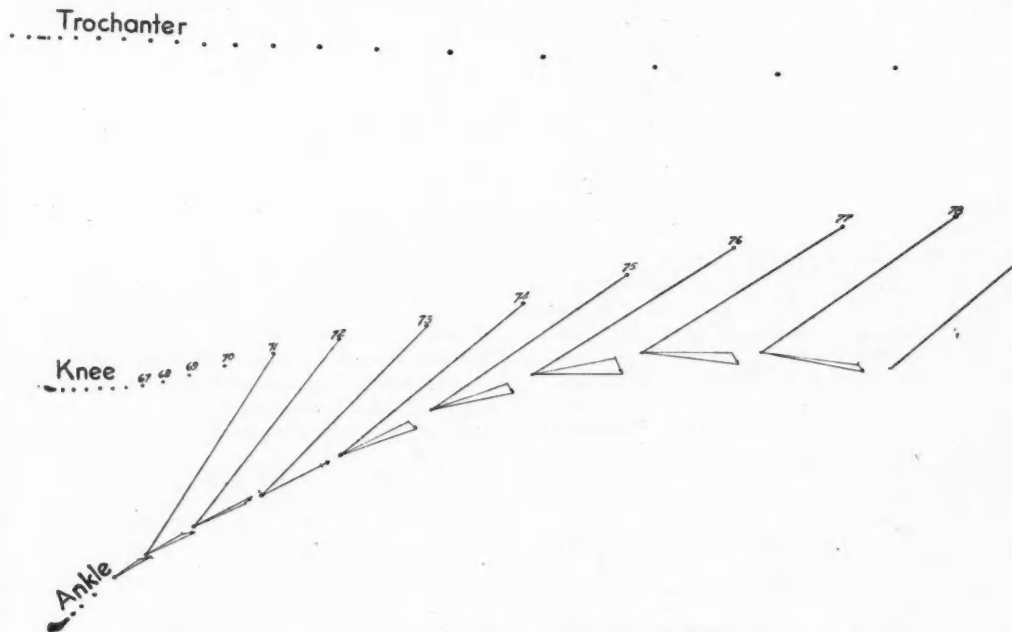


FIG. 11. Direct graphical vector evaluation of velocities and accelerations of the ankle from interrupted light experiment as in Fig. 10.

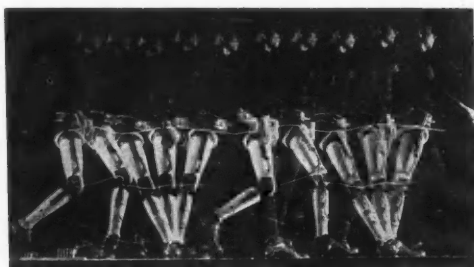


FIG. 12. Movement of an artificial limb. Compare with the normal movement shown in Fig. 10(a). [By LIFE photographer Gjon Mili; copyright TIME, Inc.]

The photograph of the subject is superimposed by a synchronous flash during his journey, set off by a delayed-action switch, tripped as he starts.

Thus successive distances between images of the lamps are proportional to the average speed during the traversal of that space interval. Therefore, average speeds for each 1/30-sec interval are directly measurable. Figure 11 is an enlargement of a portion of an original such as Fig. 10 in which the ankle, the knee and the trochanter are "spotted." The average vector acceleration can be evaluated graphically for the successive intervals, as indicated for the ankle. In this way qualitative and quantitative evaluation of the motion of parts of the limbs can be graphically portrayed. The center of mass and

mass of the lower leg may be known,¹¹ and thus an example of Newton's second law of motion may be presented for calculating the resultant force on the leg.

It is seen that the normal walking gait [Fig. 10(a)], the easiest to be assumed by a normal person, is one of approximately constant forward velocity of the center of gravity of the body (the position of the top light). There is a slight undulatory effect due to rolling on the hips as one walks. The muscles then are primarily expending energy only in accelerating and decelerating the limbs.

This example lends itself to extending qualitatively, at least, the limited kinds of motions normally studied in an elementary physics course. Too often the student is left with the idea that simple harmonic motion, or at best the motion of the planets, is the most complicated that occurs naturally.

The United States Government is financing an extensive study of normal and abnormal motions of the limbs in an effort to improve prosthetic devices, particularly for war amputees.¹⁴ Figure 12¹⁵ reveals the problem involved in improving the knee motion of the amputee using a prosthetic device.

¹⁴ Acknowledgment for Figs. 10 and 11 is gladly made to the University of California National Research Council Committee on Prosthetic Devices.

¹⁵ From *Life Magazine* (July 1, 1946), p. 92.

The notion that the scientist is professionally incapable of value judgments is one of the quaintest and most ignorant assumptions that so-called humanists can make. Nevertheless, it is generally true that, as a professional group, scientists have clung to the assumption that their primary academic business is research and the training of others to become research workers. They have not characteristically concerned themselves with academic administration except as academic administration has concerned itself with them; and they have left to what a sardonic colleague calls the "talking departments" the formulation of educational theory. The consequence is that the library of books and articles on higher education has principally been the product of nonscientific members of the faculty, and oftener the work of the humanists than of the social scientists.

In a scientific and technological age our educational theory should not be shaped mainly by non-scientists. What is equally important is that whereas able scientists make some effort to understand the social scientists and the humanists, the humanists make little effort to comprehend the scientists, so that, when the humanists attempt to discuss science in the curriculum, they tend to talk about it as if they were living in the late nineteenth century wherein a simple, mechanistic theory sufficed.—

HOWARD MUMFORD JONES, *Education and world tragedy*.

Joule's Only General Exposition of the Principle of Conservation of Energy

E. C. WATSON

California Institute of Technology, Pasadena 4, California

THE year 1847 marks an epoch in the history of science. On April 28, 1847, JAMES PRESCOTT JOULE (1818–1889) gave his first (and only) full and clear exposition of the universality of what we now call *the principle of conservation of energy*, and in June and August of the same year the significance of his careful determinations of the mechanical equivalent of heat was first recognized by his scientific peers. Because his conclusions were based upon the most careful and painstaking measurements, they have stood the test of time and remain, a hundred years later, as one of the corner stones of physical science. Many physicists will therefore wish to commemorate this centennial by re-reading JOULE's own exposition of the great and fruitful generalization he was instrumental in establishing—an exposition that was written only after he had been led by his experiments to a full realization of their general significance. The purpose of this paper is to reproduce this exposition in full so that it may be readily available. Those physicists who will also take the time to re-read OSBORNE REYNOLD's "Memoir of James Prescott Joule",¹ which has been characterized as the "best biography of a scientist in the English language," will be amply repaid.

While the Institute of France was the first national academy or institute to recognize the importance of JOULE's determinations of the mechanical equivalent of heat,² the meeting of the British Association for the Advancement of Science at Oxford on June 23, 1847, will always be memorable as the occasion on which the work on the mechanical equivalent received its first recognition by any English physicist. Fortunately both JOULE and WILLIAM THOMSON

(LORD KELVIN) have recorded what took place. Thus Joule, himself, wrote in 1885 as follows:³

It was in the year 1843 that I read a paper "On the Calorific Effects of Magneto-Electricity and the Mechanical Value of Heat" to the Chemical Section of the British Association assembled at Cork. With the exception of some eminent men, among whom I recollect with pride Dr. Apjohn, the president of the Section, the Earl of Rosse, Mr. Eaton Hodgkinson, and others, the subject did not excite much general attention; so that when I brought it forward again at the meeting in 1847, the chairman suggested that, as the business of the section pressed, I should not read my paper, but confine myself to a short verbal description of my experiments. This I endeavoured to do, and discussion not being invited, the communication would have passed without comment if a young man had not risen in the section, and by his intelligent observations created a lively interest in the new theory. The young man was William Thomson, who had two years previously passed the University of Cambridge with the highest honour, and is now probably the foremost scientific authority of the age.

KELVIN's accounts of this memorable meeting and of its consequences are fuller and more revealing. In an address⁴ delivered on the occasion of the unveiling of JOULE's statue⁵ in the Manchester Town Hall, December 7, 1893, he made this gracious statement:

I can never forget the British Association at Oxford in the year 1847, when in one of the sections I heard a paper read by a very unassuming young man who betrayed no consciousness in his manner that he had a great idea to unfold. I was tremendously struck with the paper. I at first thought it could not be true because it was different from Carnot's theory, and immediately after the reading of the paper I had a few words of conversation with the author James Joule, which was the beginning of our forty years' acquaintance and friendship. On the evening of the same day that very valuable Institution of the British Association, its conversazione, gave us opportunity for a good hour's talk and discussion over all that

¹ O. Reynold, *Mem. Proc. Manchester Lit. and Phil. Soc.* (4) 6, 1–191 (1892).

² "Expériences sur l'identité entre le calorique et la force mécanique. Détermination de l'équivalent par la chaleur dégagée pendant la friction du mercure. Par M. J. P. Joule," *Compt. Rend.*, Aug. 23, 1847. In 1881, Joule added the following note: "The commissioners were Biot, Pouillet, and Regnault. I had the honour to present the iron vessel with its revolving paddle-wheel to the last named eminent physicist" [*The scientific papers of James Prescott Joule* (London, 1884), p. 283].

³ *Joint scientific papers of James Prescott Joule* (London, 1887), p. 215. See also *Joule and the study of energy*, by Alex Wood (Bell, 1925), pp. 51–52.

⁴ *Nature* 49, 164 (1893); *Popular lectures and addresses* (Macmillan, 1894), vol. 2, pp. 566–568; A. Wood, *Joule and the study of energy* (Bell, 1925), pp. 54–56.

⁵ See *Am. J. Physics*, 9, 111 (1941) for a reproduction of a photograph of this statue.

either of us knew of thermodynamics. I gained ideas which had never entered my mind before, and I thought I too suggested something worthy of Joule's consideration when I told him of Carnot's theory. Then and there in the Radcliffe Library, Oxford, we parted, both of us, I am sure, feeling that we had much more to say to one another and much matter for reflection in what we had talked over that evening. But what was my surprise a fortnight later when, walking down the valley of Chamounix, I saw in the distance a young man walking up the road toward me and carrying in his hand something that looked like a stick, but which he was using neither as an Alpenstock nor as a walking stick. It was Joule with a long thermometer in his hand, which he would not trust by itself in the *char-à-banc* coming slowly up the hill behind him lest it should get broken. But there comfortably and safely seated on the *char-à-banc* was his bride—the sympathetic companion and sharer in his work of after years. He had not told in Section A or in the Radcliffe Library that he was going to be married in three days, but now in the valley of Chamounix, he introduced me to his young wife. We appointed to meet again a fortnight later at Martigny to make experiments on the heat of a waterfall (Sallanches) with that thermometer: and afterwards we met again and again and again, and from that time indeed remained close friends till the end of Joule's life. I had the great pleasure and satisfaction for many years, beginning just forty years ago, of making experiments along with Joule which led to some important results in respect to the theory of thermodynamics. This is one of the most valuable recollections of my life, and is indeed as valuable a recollection as I can conceive in the possession of any man interested in science.

Much the same story is told, but with some additional details, in a letter to J. T. BOTTOMLEY which forms a part of a brief biographical memoir.⁶ Only such parts are quoted here as help to fill out the picture just sketched:

I made Joule's acquaintance at the Oxford Meeting, and it quickly ripened into a life-long friendship. I heard his paper read at the sections, and felt strongly impelled to rise and say that it must be wrong, because the true mechanical value of heat given, suppose to warm water, must, for small differences of temperature, be proportional to the square of its quantity. I knew from Carnot's law that this must be true (and it is true; only now I call it "motivity," in order not to clash with Joule's "Mechanical Value"). But as I listened on and on, I saw that (though Carnot had a vitally important truth not to be abandoned) Joule had certainly a great truth and a great discovery, and a most important measurement to bring forward. So instead of rising with my objection to the meeting,

I waited till it was over and said my say to Joule himself at the end of the meeting. This made my first introduction to him. . . .

Joule's paper at the Oxford meeting made a great sensation. Faraday was there, and was much struck with it, but did not enter fully into the new views. It was many years after that, before any of the scientific chiefs began to give their adhesion. It was not long after when Stokes told me he was inclined to be a Jouleite.

Miller or Graham, or both, were for many years quite incredulous as to Joule's results, because they all depended on fractions of a degree of temperature—sometimes very small fractions. His boldness in making such large conclusions from such very small observational effects, is almost as noteworthy and admirable as his skill in extorting accuracy from them. I remember distinctly at the Royal Society, I think it was either Graham or Miller saying simply he did not believe Joule because he had nothing but hundredths of a degree to prove his case by.

JOULE's only exposition of the principle of conservation of energy in all its generality was given in a popular lecture at St. Ann's Church Reading Room in Manchester on April 28, 1847. This lecture, entitled "On Matter, Living Force, and Heat,"⁷ is reproduced in full at the end of this paper. In order to appreciate it properly it is necessary to remember that the terms "potential energy" and "kinetic energy," as we now use them in scientific parlance, had not yet come into existence. The term "potential energy" was first used by the Scottish engineer, W. J. M. RANKINE, in a paper⁸ read before the Philosophical Society of Glasgow in 1853, and the term "kinetic energy" was introduced by LORD KELVIN in 1879⁹ or a little earlier.¹⁰

It is interesting to compare this exposition of JOULE's with that of JULIUS ROBERT MAYER (1814–1878) published in 1842,¹¹ the first published paper to contain a clear statement of the law of conservation of energy. MAYER's account, although published first, was largely intuitive¹²

⁷ *The scientific papers of James Prescott Joule* (London, 1884), pp. 265–276.

⁸ *Scientific papers* (Griffin, 1881), pp. 203, 229.

⁹ *Treatise on natural philosophy*.

¹⁰ *Mathematical and physical papers*, vol. 2, p. 34, footnote. ¹¹ "Bemerkungen über die Kräfte der unbelebten Natur," *Ann. Chem. Pharm.* 42, 233 (1842); a translation by G. C. Foster, first published in *Phil. Mag.* 24, 371 (1862) is reproduced in W. F. Magie's *A source book in physics* (McGraw-Hill, 1935), pp. 197–203, and *Isis* 13, 27–33 (1929).

¹² For an objective appraisal of the merits of the various claims put forward for Mayer and Joule see G. Sarton, *Isis* 13, 18 (1929).

⁶ *Nature* 26, 617 (1882).

while JOULE's was based upon "measurement, rigorous experiment and observation, and philosophic thought all round the field of physical science."¹³ His experiments on the mechanical equivalent covered a period of about 40 years. Reference to the original papers themselves is necessary if one is to form a correct idea of the enormous experimental labor they represent.¹⁴

As J. CLERK MAXWELL wrote to BALFOUR STEWART,¹⁵ "there are only a very few men who have stood in a similar position and who have been urged by the love of some truth which they were confident was to be found though its form was as yet undefined to devote themselves to minute observations and patient manual and mental toil in order to bring their thoughts into exact accordance with things as they are."

ON MATTER, LIVING FORCE, AND HEAT

By

J. P. Joule,

Secretary of the Manchester Literary and Philosophical Society

In our notion of matter two ideas are generally included, namely those of *impenetrability* and *extension*. By the extension of matter we mean the space which it occupies; by its impenetrability we mean that two bodies cannot exist at the same time in the same place. Impenetrability and extension cannot with much propriety be reckoned among the *properties* of matter, but deserve rather to be called its *definitions*, because nothing that does not possess the two qualities bears the name of matter. If we conceive of impenetrability and extension we have the idea of matter, and of matter only.

Matter is endowed with an exceedingly great variety of wonderful properties, some of which are common to all matter, while others are present variously, so as to constitute a difference between one body and another. Of the first of these classes, the attraction of gravitation is one of the most important. We observe its presence readily in all solid bodies, the component parts of which are, in the opinion of Majocchi, held together by this force. If we break the body in pieces, and remove the separate pieces to a distance from each other, they will still be found to attract each other, though in a very slight degree, owing to the force being one which diminishes very rapidly as the bodies are removed further from one another. The larger the bodies are the more powerful is the force of attraction subsisting between them. Hence, although the force of attraction between small bodies can only be appreciated by the most delicate apparatus except in the case of contact, that which is occasioned by a body of immense magnitude, such as the earth, becomes very considerable. This attraction of bodies towards the earth constitutes what is called their *weight* or *gravity*, and is always exactly proportional to the quantity of matter. Hence, if any body be found to weigh 2 lb., while another only weighs 1 lb., the former will contain exactly twice as much matter as the latter; and this is the case, whatever the bulk of the bodies may be: 2-lb. weight of air contains exactly twice the quantity of matter that 1 lb. of lead does.

Matter is sometimes endowed with other kinds of attraction besides the attraction of gravitation; sometimes also it possesses the faculty of *repulsion*, by which force the particles tend to separate further from each other. Wherever these forces exist, they do not supersede the attraction of gravitation. Thus the weight of a piece of iron or steel is in no way affected by imparting to it the magnetic virtue.

Besides the force of gravitation, there is another very remarkable property displayed in an equal degree by every kind of matter—its perseverance in any condition, whether of rest or motion, in which it may have been placed. This faculty has received the name of *inertia*, signifying passiveness, or the inability of any thing to change its own state. It is in consequence of this property that a body at rest cannot be set in motion without the application of a certain amount of force to it, and also that when once the body has been set in motion it will never stop of itself, but continue to move straight forwards with a uniform velocity until acted upon by another force, which, if applied contrary to the direction of motion, will retard it, if in the same direction will accelerate it, and if sideways will cause it to move in a curved direction. In the case in which the force is applied contrary in direction, but equal in degree to that which set the body first in motion, it will be entirely deprived of motion whatever time may have elapsed since the first impulse, and to whatever distance the body may have travelled.

From these facts it is obvious that the force expended in setting a body in motion is carried by the body itself, and exists with it and in it, throughout the whole course of its motion. This force possessed by moving bodies is termed by mechanical philosophers *vis viva*, or *living force*. The term may be deemed by some inappropriate, inasmuch as there is no life, properly speaking, in question; but it is

¹³ Lord Kelvin, *Popular lectures and addresses* (Macmillan, 1894), vol. 2, p. 564.

¹⁴ See *The scientific papers of James Prescott Joule* London, (1884, 1887).

¹⁵ Letter preserved among the historical items in the Department of Pure and Applied Physics, College of Technology, Manchester, England. See "The Joule collection in the College of Technology, Manchester," by H. Lowery, *J. Sci. Inst.* 7, 369 (1930); 8, 1 (1931).



JAMES PRESCOTT JOULE (1818-1889). [From an engraving by H. Manesse after the painting by George Patten in the possession of the Manchester Literary and Philosophical Society.]

useful in order to distinguish the moving force from that which is stationary in its character, as the force of gravity. When, therefore, in the subsequent parts of this lecture I employ the term *living force*, you will understand that I simply mean the force of bodies in motion. The living force of bodies is regulated by their weight and by the velocity of their motion. You will readily understand that if a body of a certain weight possess a certain quantity of living force, twice as much living force will be possessed

by a body of twice the weight, provided both bodies move with equal velocity. But the law by which the *velocity* of a body regulates its living force is not so obvious. At first sight one would imagine that the living force would be simply proportional to the velocity, so that if a body moved twice as fast as another, it would have twice the impetus or living force. Such, however, is not the case; for if three bodies of equal weight move with the respective velocities of 1, 2, and 3 miles per hour, their living forces will be

found to be proportional to those numbers multiplied by themselves, *viz.* to 1×1 , 2×2 , 3×3 , or 1, 4, and 9, the squares of 1, 2, and 3. This remarkable law may be proved in several ways. A bullet fired from a gun at a certain velocity will pierce a block of wood to only one quarter of the depth it would if propelled at twice the velocity. Again, if a cannon-ball were found to fly at a certain velocity when propelled by a given charge of gun-powder, and it were required to load the cannon so as to propel the ball with twice that velocity, it would be found necessary to employ four times the weight of powder previously used. Thus, also, it will be found that a railway-train going at 70 miles per hour possesses 100 times the impetus, or living force, that it does when travelling at 7 miles per hour.

A body may be endowed with living force in several ways. It may receive it by the impact of another body. Thus, if a perfectly elastic ball be made to strike another similar ball of equal weight at rest, the striking ball will communicate the whole of its living force to the ball struck, and, remaining at rest itself, will cause the other ball to move in the same direction and with the same velocity that it did itself before the collision. Here we see an instance of the facility with which living force may be transferred from one body to another. A body may also be endowed with living force by means of the action of gravitation upon it through a certain distance. If I hold a ball at a certain height and drop it, it will have acquired when it arrives at the ground a degree of living force proportional to its weight and the height from which it has fallen. We see, then, that living force may be produced by the action of gravity through a given distance or space. We may, therefore, say that the former is of equal value, or *equivalent*, to the latter. Hence, if I raise a weight of 1 lb. to the height of one foot, so that gravity may act on it through that distance, I shall communicate to it that which is of equal value or equivalent to a certain amount of living force; if I raise the weight to twice the height, I shall communicate to it the equivalent of twice the quantity of living force. Hence, also, when we compress a spring, we communicate to it the equivalent to a certain amount of living force; for in that case we produce molecular attraction between the particles of the spring through the distance they are forced asunder, which is strictly analogous to the production of the attraction of gravitation through a certain distance.

You will at once perceive that the living force of which we have been speaking is one of the most important qualities with which matter can be endowed, and, as such, that it would be absurd to suppose that it can be destroyed, or even lessened, without producing the equivalent of attraction through a given distance of which we have been speaking. You will therefore be surprised to hear that until very recently the universal opinion has been that living force could be absolutely and irrevocably destroyed at any one's option. Thus, when a weight falls to the ground, it has been generally supposed that its living force is absolutely annihilated, and that the labour which may have been expended in raising it to the elevation from which it fell has been entirely thrown away and wasted

without the production of any permanent effect whatever. We might reason, *à priori*, that such absolute destruction of living force cannot possibly take place, because it is manifestly absurd to suppose that the powers with which God has endowed matter can be destroyed any more than that they can be created by man's agency; but we are not left with this argument alone, decisive as it must be to every unprejudiced mind. The common experience of every one teaches him that living force is not *destroyed* by the friction or collision of bodies. We have reason to believe that the manifestations of living force on our globe are, at the present time, as extensive as those which have existed at any time since its creation, or, at any rate, since the deluge—that the winds blow as strongly, and the torrents flow with equal impetuosity now, as at the remote period of 4000 or even 6000 years ago; and yet we are certain that, through that vast interval of time, the motions of the air and of the water have been incessantly obstructed and hindered by friction. We may conclude, then, with certainty, that these motions of air and water, constituting living force, are not *annihilated* by friction. We lose sight of them, indeed, for a time; but we find them again reproduced. Were it not so, it is perfectly obvious that long ere this all nature would have come to a dead standstill. What, then, may we inquire, is the cause of this apparent anomaly? How comes it to pass that, though in almost all natural phenomena we witness the arrest of motion and the apparent destruction of living force, we find that no waste or loss of living force has actually occurred? Experiment has enabled us to answer these questions in a satisfactory manner; for it has shown that, wherever living force is *apparently* destroyed, an equivalent is produced which in process of time may be reconverted into living force. This equivalent is *heat*. Experiment has shown that wherever living force is apparently destroyed or absorbed, heat is produced. The most frequent way in which living force is thus converted into heat is by means of friction. Wood rubbed against wood or against any hard body, metal rubbed against metal or against any other body—in short, all bodies, solid or even liquid, rubbed against each other are invariably heated, sometimes even so far as to become red-hot. In all these instances the quantity of heat produced is invariably in proportion to the exertion employed in rubbing the bodies together—that is, to the living force absorbed. By fifteen or twenty smart and quick strokes of a hammer on the end of an iron rod of about a quarter of an inch in diameter placed upon an anvil an expert blacksmith will render that end of the iron visibly red-hot. Here heat is produced by the absorption of the living force of the descending hammer in the soft iron; which is proved to be the case from the fact that the iron cannot be heated if it be rendered hard and elastic, so as to transfer the living force of the hammer to the anvil.

The general rule, then, is, that wherever living force is *apparently* destroyed, whether by percussion, friction, or any similar means, an exact equivalent of heat is restored. The converse of this proposition is also true, namely, that heat cannot be lessened or absorbed without the production of living force, or its equivalent attraction through space. Thus, for instance, in the steam-engine it will be found that

the power gained is at the expense of the heat of the fire—that is, that the heat occasioned by the combustion of the coal would have been greater had a part of it not been absorbed in producing and maintaining the living force of the machinery. It is right, however, to observe that this has not as yet been demonstrated by experiment. But there is no room to doubt that experiment would prove the correctness of what I have said; for I have myself proved that a conversion of heat into living force takes place in the expansion of air, which is analogous to the expansion of steam in the cylinder of the steam-engine. But the most convincing proof of the conversion of heat into living force has been derived from my experiments with the electromagnetic engine, a machine composed of magnets and bars of iron set in motion by an electrical battery. I have proved by actual experiment that, in exact proportion to the force with which this machine works, heat is abstracted from the electrical battery. You see, therefore, that living force may be converted into heat, and that heat may be converted into living force, or its equivalent attraction through space. All three, therefore—namely, heat, living force, and attraction through space (to which I might also add *light*, were it consistent with the scope of the present lecture)—are mutually convertible into one another. In these conversions nothing is ever lost. The same quantity of heat will always be converted into the same quantity of living force. We can therefore express the equivalency in definite language applicable at all times and under all circumstances. Thus the attraction of 817 lb. through the space of one foot is equivalent to, and convertible into, the living force possessed by a body of the same weight of 817 lb. when moving with the velocity of eight feet per second, and this living force is again convertible into the quantity of heat which can increase the temperature of one pound of water by one degree Fahrenheit. The knowledge of the equivalency of heat to mechanical power is of great value in solving a great number of interesting and important questions. In the case of the steam-engine, by ascertaining the quantity of heat produced by the combustion of coal, we can find out how much of it is converted into mechanical power, and thus come to a conclusion how far the steam-engine is susceptible of further improvements. Calculations made upon this principle have shown that at least ten times as much power might be produced as is now obtained by the combustion of coal. Another interesting conclusion is, that the animal frame, though destined to fulfil so many other ends, is as a machine more perfect than the best contrived steam-engine—that is, is capable of more work with the same expenditure of fuel.

Behold, then, the wonderful arrangements of creation. The earth in its rapid motion round the sun possesses a degree of living force so vast that, if turned into the equivalent of heat, its temperature would be rendered at least 1000 times greater than that of red-hot iron, and the globe on which we tread would in all probability be rendered equal in brightness to the sun itself. And it cannot be doubted that if the course of the earth were changed so that it might fall into the sun, that body, so far from being cooled down by the contact of a comparatively cold body,

would actually blaze more brightly than before in consequence of the living force with which the earth struck the sun being converted into its equivalent of heat. Here we see that our existence depends upon the *maintenance* of the living force of the earth. On the other hand, our safety equally depends in some instances upon the *conversion* of living force into heat. You have, no doubt, frequently observed what are called *shooting-stars*, as they appear to emerge from the dark sky of night, pursue a short and rapid course, burst, and are dissipated in shining fragments. From the velocity with which these bodies travel, there can be little doubt that they are small planets which, in the course of their revolution round the sun, are attracted and drawn to the earth. Reflect for a moment on the consequences which would ensue, if a hard meteoric stone were to strike the room in which we are assembled with a velocity sixty times as great as that of a cannon-ball. The dire effects of such a collision are effectually prevented by the atmosphere surrounding our globe, by which the velocity of the meteoric stone is checked and its living force converted into heat, which at last becomes so intense as to melt the body and dissipate it into fragments too small probably to be noticed in their fall to the ground. Hence it is that, although multitudes of shooting-stars appear every night, few meteoric stones have been found, those few corroborating the truth of our hypothesis by the marks of intense heat which they bear on their surfaces.

Descending from the planetary space and firmament to the surface of our earth, we find a vast variety of phenomena connected with the conversion of living force and heat into one another, which speak in language which cannot be misunderstood of the wisdom and beneficence of the Great Architect of nature. The motion of air which we call *wind* arises chiefly from the intense heat of the torrid zone compared with the temperature of the temperate and frigid zones. Here we have an instance of heat being converted into the living force of currents of air. These currents of air, in their progress across the sea, lift up its waves and propel the ships; whilst in passing across the land they shake the trees and disturb every blade of grass. The waves by their violent motion, the ships by their passage through a resisting medium, and the trees by the rubbing of their branches together and the friction of their leaves against themselves and the air, each and all of them generate heat equivalent to the diminution of the living force of the air which they occasion. The heat thus restored may again contribute to raise fresh currents of air; and thus the phenomena may be repeated in endless succession and variety.

When we consider our own animal frames, "fearfully and wonderfully made," we observe in the motion of our limbs a continual conversion of heat into living force, which may be either converted back again into heat or employed in producing an attraction through space, as when a man ascends a mountain. Indeed the phenomena of nature, whether mechanical, chemical, or vital, consist almost entirely in a continual conversion of attraction through space, living force, and heat into one another. Thus it is that order is maintained in the universe—nothing is

deranged, nothing ever lost, but the entire machinery, complicated as it is, works smoothly and harmoniously. And though, as in the awful vision of Ezekiel, "wheel may be in the middle of wheel," and every thing may appear complicated and involved in the apparent confusion and intricacy of an almost endless variety of causes, effects, conversions, and arrangements, yet is the most perfect regularity preserved—the whole being governed by the sovereign will of God.

A few words may be said, in conclusion, with respect to the real nature of heat. The most prevalent opinion, until of late, has been that it is a *substance* possessing, like all other matter, impenetrability and extension. We have, however, shown that heat can be converted into living force and into attraction through space. It is perfectly clear, therefore, that unless matter can be converted into attraction through space, which is too absurd an idea to be entertained for a moment, the hypothesis of heat being a substance must fall to the ground. Heat must therefore consist of either living force or of attraction through space. In the former case we can conceive the constituent particles of heated bodies to be, either in whole or in part, in a state of motion. In the latter we may suppose the particles to be removed by the process of heating, so as to exert attraction through greater space. I am inclined to believe that both of these hypotheses will be found to hold good—that in some instances, particularly in the case of *sensible* heat, or such as is indicated by the thermometer, heat will be found to consist in the living force of the particles of the bodies in which it is induced; whilst in others, particularly in the case of *latent* heat, the phenomena are produced by the separation of particle from particle, so as to cause them to attract one another through a greater space. We may conceive, then, that the communication of heat to a body consists, in fact, in the communication of impetus, or living force, to its particles. It will perhaps appear to some of you something strange that a body apparently quiescent should in reality be the seat of motions of great rapidity; but you will observe that the bodies themselves, considered as wholes, are not supposed to be in motion. The constituent particles, or atoms of the bodies, are supposed to be in motion, without producing a gross motion of the whole mass. These particles, or atoms, being far too small to be seen even by the help of the most powerful microscopes, it is no wonder that we cannot observe their motion. There is therefore reason to suppose that the particles of all bodies, their constituent atoms, are in a state of motion almost too rapid for us to conceive, for the phenomena cannot be otherwise explained. The velocity of the atoms of water, for instance, is at least equal to a mile per second of time. If, as there is reason to think, some particles are at rest while others are in motion, the velocity of the latter will be proportionally greater. An increase of the velocity of revolution of the particles will constitute an increase of temperature, which may be distributed among the neighbouring bodies by what is called *conduction*—that is, on the present hypothesis, by the communication of the increased motion from the particles of one body to those of another. The velocity of the particles being further

increased, they will tend to fly from each other in consequence of the centrifugal force overcoming the attraction subsisting between them. This removal of the particles from each other will constitute a new condition of the body—it will enter into the state of fusion, or become melted. But, from what we have already stated, you will perceive that, in order to remove the particles violently attracting one another asunder, the expenditure of a certain amount of living force or heat will be required. Hence it is that heat is always absorbed when the state of a body is changed from solid to liquid, or from liquid to gas. Take, for example, a block of ice cooled down to zero; apply heat to it, and it will gradually arrive at 32° , which is the number conventionally employed to represent the temperature at which ice begins to melt. If, when the ice has arrived at this temperature, you continue to apply heat to it, it will become melted; but its temperature will not increase beyond 32° until the whole has been converted into water. The explanation of these facts is clear on our hypothesis. Until the ice has arrived at the temperature of 32° the application of heat increases the velocity of rotation of its constituent particles; but the instant it arrives at that point, the velocity produces such an increase of the centrifugal force of the particles that they are compelled to separate from each other. It is in effecting this separation of particles strongly attracting one another that the heat applied is *then* spent; not in increasing the velocity of the particles. As soon, however, as the separation has been effected, and the fluid water produced, a further application of heat will cause a further increase of the velocity of the particles, constituting an increase of temperature, on which the thermometer will immediately rise above 32° . When the water has been raised to the temperature of 212° , or the boiling-point, a similar phenomenon will be repeated; for it will be found impossible to increase the temperature beyond that point, because the heat then applied is employed in separating the particles of water so as to form steam, and not in increasing their velocity and living force. When, again, by the application of cold we condense the steam into water, and by a further abstraction of heat we bring the water to the solid condition of ice, we witness the repetition of similar phenomena in the reverse order. The particles of steam, in assuming the condition of water, fall together through a certain space. The living force thus produced becomes converted into heat, which must be removed before any more steam can be converted into water. Hence it is always necessary to abstract a great quantity of heat in order to convert steam into water, although the temperature will all the while remain exactly at 212° ; but the instant that all the steam has been condensed, the further abstraction of heat will cause a diminution of temperature, since it can only be employed in diminishing the velocity of revolution of the atoms of water. What has been said with regard to the condensation of steam will apply equally well to the congelation of water.

I might proceed to apply the theory to the phenomena of combustion, the heat of which consists in the living force occasioned by the powerful attraction through space

of the combustible for the oxygen, and to a variety of other thermo-chemical phenomena; but you will doubtless be able to pursue the subject further at your leisure.

I do assure you that the principles which I have very

imperfectly advocated this evening may be applied very extensively in elucidating many of the abstruse as well as the simple points of science, and that patient inquiry on these grounds can hardly fail to be amply rewarded.

Systems of Electrical Units

JOHN A. ELDRIDGE

The State University of Iowa, Iowa City, Iowa

IN his article on "Physical units and standards" in Eshbach's *Handbook of engineering fundamentals* (Wiley), E. Weber says that a "complete dimension system in mechanics must have three, in thermodynamics four, and in electromagnetism four fundamental dimensions." Yet elsewhere in this article he says: "The physical quantities, arbitrarily chosen to define new quantities or derived quantities, are called fundamental physical quantities. Their number may vary according to needs and convenience. Physical quantities which appear to be fundamental in some one special field may be derived quantities in some other field."

Ultimately each physical quantity—length, area, mass, force, and so forth—is essentially different; the expression of one in terms of another is a convention of science. For example, consider the geometric quantities area and length. These are essentially unlike quantities; area is the capacity of a surface for trees or chairs or square centimeters; length is something entirely different. Both are intuitive concepts with no self-evident relation between them. The unit of area is some arbitrarily chosen surface, say a tree-spread or (rather better for scientific purposes) an acre; the unit of length is an arbitrarily chosen length, say a centimeter. The magnitude of the area (A) means the number of area units and that of the length (l) means the number of length units. Now it can be shown that for similar figures the area (in any units) is proportional to the square of some characteristic length. For example, for circles with radii r ,

$$A = kr^2, \quad (1)$$

where k is a constant of proportionality with dimensions (area/length²). We must call k a dimensional constant because its value depends on the particular units used. In fact, it depends only on the units, and we call it a conversion factor instead of a geometric variable. We may now decide to adopt the universal practice of always using units that will give k some convenient fixed value. For example, we may adopt the relation

$$A = 1 \cdot r^2. \quad (2)$$

We then say that "the proportionality constant is dimensionless," that the equation is a "definition of area" (meaning a definition of its magnitude, the numeric A , as we agree to use it). If the unit of length is a centimeter we call the unit of area a *circular centimeter* (analogous to a

circular 'mil'). We can then develop geometric equations that give correct area magnitudes as here defined. We are more likely, however, to adopt a different "definition of area," making $k = \pi$, which makes $A = l^2$ for a square; we call the area unit a *square centimeter*. Observe that introducing such a convention as this has a certain advantage and a certain disadvantage; it simplifies the equation and gives us one less constant to look up in tables. On the other hand, it restricts us to the use of consistent units; we cannot find the area in acres directly from Eq. (1), as we can from Eq. (2).

* * *

In a consistent system of units the number of dimensions that are regarded as fundamental is arbitrary. In mechanics it is conventional to consider length, mass and time as fundamental. In the equation $F = kma$, k is regarded as dimensionless, is given a definite value (usually unity), and the equation is called a defining equation for F ; thus F is not regarded as a fundamental dimension. On the other hand, in the equation for gravitational attraction, $F = Gm_1m_2/r^2$, G is not regarded as dimensionless; this equation is not regarded as a defining equation of mass. Mass is regarded as a fundamental dimension. This is entirely arbitrary; if we have three fundamental dimensions in mechanics it is only as a matter of convenience and convention.

In electricity the convention has not been so clearly established. Coulomb's law,

$$F = q_1q_2/kr^2, \quad (3)$$

for the attraction between point charges in empty space may be regarded as a definition of the charge magnitude q , k being given some definite value, usually 1. If we adopt this convention the defining equation is

$$F = q_1q_2/r^2. \quad (4)$$

Electrical quantities are then measured in terms of mechanical quantities. We may consistently have an fps system or an mks system or a cgs system of electrical units. The cgs unit of charge is the statcoulomb. Unless the definition of q [Eq. (4)] is changed, there can be no other cgs unit of charge, such as the abcoulomb.

On the other hand, electric charge may be regarded as a fourth fundamental dimension, its unit being unrelated to the mechanical units of the system. The constant k then has the dimensions $[Q^2T^2/L^3M]$. In this four-dimensional systematology, the meter-kilogram-second-coulomb (mksc) system is most generally recognized.¹ In physics the centimeter-gram-second-statcoulomb (cgss) and the centimeter-gram-second-abcoulomb (cgsa) systems are also used.

As between the three-dimensional and the four-dimensional systematologies there has always been some difference of opinion as to which really was the more "convenient and conventional." (There has also been considerable meaningless argument as to the logical necessity of the one system or the other.) Gaussian units (three-dimensional) have usually seemed convenient to physicists. Whatever point of view we take, the charge units commonly used in theoretical physics (statcoulomb, abcoulomb) actually are based on mechanical quantities and were adopted to make the proportionality constant unity in certain simple equations. The three-dimensionalists wish to exploit this simplicity by dropping the constant. On the other hand, the engineer has found it convenient to use other units. In particular he wishes to use practical units (coulomb, volt, and so forth). The physicist uses these same practical units, but usually regards them as secondary and reduces them to "fundamental units" in applying them to mechanical problems. The practical units are related to the fundamental units by powers of 10 to facilitate this reduction. But the engineer working with these and other units wishes to have his formulas directly applicable without reduction. He finds it convenient to express the reduction factor in the formula (quantity k).

¹ The fundamental electrical unit is often given as the ohm (mks Ω units), but it is immaterial which quantity is chosen as fundamental, and the various systems are more easily compared if we use the coulomb.

It is not our purpose to discuss the relative merits of the two systematologies but to describe them. What may be "convenient and conventional" in one domain of physics may not be so in another. But the advocates of different systems are likely to misunderstand one another. An engineer is likely to consider the physicist who confuses oersteds $[I/L]$ and gaussses $[M/IT^2]$ as logically inconsistent; the physicist is likely to consider the distinction as trivial since the two quantities are numerically the same;² they are also dimensionally the same if we accept Eq. (4) as a definition.

It is generally agreed that all other electrical quantities (including magnetic quantities) can be defined in terms of charge, together with mechanical concepts. But now there is another difference in usage: two distinct systems of definition are used for magnetic quantities. In theoretical physics magnetic induction B is likely to be defined by the equation for the force on a wire: $F = (i/c)lB$; this is the "Gaussian definition." More commonly we find it defined by the equation³ $F = ilB$. The consistent unit of B in the cgs (three-dimensional) system or in the cgss (electrostatic) system is the gauss if we use the first definition; the unit of B (in these systems) is 3×10^{10} gauss cm/sec if we use the second definition. Other units are defined in terms of B , and the constant c , if expressed here, occurs in the definitions of all magnetic quantities.

This is a serious disagreement. These alternatively defined quantities differ not simply in magnitude, as between a "rational" and an "irrational" system, but differ dimensionally from one another, since c is a velocity.⁴ Since

² This does not mean that the B -field and the H -field are numerically equal in a magnetic medium; they differ because of their difference in structure. But the sourceless field (which we shall always call a B -field) has the same numerical value when it is measured in gaussses as when it is measured in oersteds (our B' -field), as is often necessary.

³ It is more commonly defined in terms of electromagnetic induction, and this equation is regarded as a "law." The two points of view are equivalent, and this one is more convenient for our presentation.

⁴ It is sometimes maintained that c is dimensionless; see the article by Weber in Eshbach's *Handbook of Engineering Fundamentals*. As we use it, c is the velocity of light; v/c is dimensionless; c^2 is the ratio of the constants of proportionality in the fundamental laws (a) and (i) in Table I. On the other hand, the ratio between the abcoulomb and the statcoulomb, two charge units commonly used in four-dimensional systems, is equal to charge units the value of c in cgs units. This is best regarded as a dimensionless

the same symbol is used in different senses in the two systems, we have two sets of magnetic equations—one in which c is expressed and one in which it is included in the magnetic quantity.

TABLE I. Comparison of cgs and mksc units.

Gauss (cgs) system		Giorgi (mksc) system	
		Covariant quantities	Contravariant quantities
Coulomb law (free space)			
(a) $F = q_1 q_2 / r^2$, defines q		$F = q^2 / r^2$, defines k^*	
Repulsion in homogeneous fluid (defines ϵ)			
(b) $F = q_1 q_2 / \epsilon r^2$		$F = q^2 / \epsilon r^2$	
Electric current			
(c) $i = q/t$ (statamp)		$i = q/t$ (amp)	
Electric potential			
(d) $V = W/q = q'/\epsilon r$ (statvolt)		$V = W/q'$ (volt)	
Field strength ("irrotational")			
(e) $E = F/q = q'/\epsilon r^2$ $= Q/r^2$		$E = F/q'$ (newton/coul, or volt/m)	
Electric displacement ("sourceless" field)			
(f) $D = q/r^2 = \epsilon F/q$ $= F/Q$		$D = F/Q'$ (coul/m ²) ($D = F/Q' = D'/k^*$, same unit as E .)	
Electric moment of charge pair of separation L			
(g) Moment $= qL = T/E^*$		Moment $= q'L = T/E$, (coul/m) ²	
Electric moment M_e , polarization P_e , pole strength m_e of electret			
(h) $M_e = T/E^*$ $P_e = M_e/\text{volume}$ $m_e = M_e/\text{length}$		$M_e = T/E$ $P_e = M_e/\text{volume}$ $m_e = M_e/\text{length}$	
Ampère's law (parallel wires, free space)			
(i) $F = \frac{2i_1 i_2}{c^2 r}$		$F = \frac{2i_1 i_2}{k^* c^2 r} \frac{K \cdot 2i_1 i_2}{r}$, $K = \frac{1}{k^* c^2}$	
Attraction in homogeneous fluid (defines μ)			
(j) $F = \mu 2i_1 i_2 / c^2 r$		$F = K \cdot \mu 2i_1 i_2 / r$	
Reduced current ("current in electromagnetic units")			
(k) $i = i/c$			
"Irrotational" field (H)			
(l) $H = 2i/r = F/(I/c)$ (gauss)		$H = 2i/r$ (amp/m) (cgaa: oersted) ($H = F/\mu i^*$ or F/I' , same unit as B .)	
"Sourceless" field (B)			
(m) $B = 2\mu i/r$ or $2(I/c)/r = F/i$ (gauss)		$B = 2\mu i/r$ or $2I/r$ (weber/m ²) (cgaa: gauss)	
Magnetic moment of solenoid of area A			
(n) Moment $= NI A = T/B^*$		Moment $= N' A = T/B$	
Magnetic moment M_m , intensity of magnetization P_m , pole strength m_m of permanent magnet			
(o) $M_m = T/B^*$ $P_m = M_m/\text{volume}$ $m_m = M_m/\text{length}$		$M_m = T/B$ $P_m = M_m/\text{volume}$ $m_m = M_m/\text{length}$ } Straton's usage	

* Torque T with axis perpendicular to field.

number. It would actually be dimensionless (in the most literal sense) only if the two charge units were defined in terms of each other, without introduction of mechanical concepts. So long as the statcoulomb and abcoulomb are defined to make the constants k and K unity they still have (in the literal sense) mechanical dimensions; their ratio depends on the accepted values of the meter and the second and is equal to the velocity of light in centimeters per second. This, however, is trivial. Our constant c depends on the system of units used; the number of statcoulombs in an abcoulomb does not.

It is not maintained that this difference in definition need cause confusion in actual practice, because it need not. But one hunting for logical inconsistency can find it when he finds the relation between the "same" quantities expressed in two inconsistent equations. The engineer accustomed to the second definition sometimes refers to the Gaussian system of units (using statcoulombs and gauss) as a "hybrid system."

The equations with c suppressed are formally simpler and so are generally used. But there is no absolute necessity about the second definition. Indeed, there are a number of reasons for preferring the first (Gaussian) definition in theoretical physics:

(i) It can be maintained that since magnetism is essentially a relativistic effect, the velocity limit c should be expressed, as in the Gaussian equations. Expressed in terms of electron-charge linear density η , the force on a wire (perpendicular to the field) is $F = \eta(v/c)IB$; the attraction between two long parallel wires, separated by a distance r and carrying equal currents ηv , is

$$F = \frac{2\eta^2 l}{r} \left(\frac{v}{c} \right)^2, \quad (5)$$

as compared with the force between charged wires,

$$F = 2\eta^2 l / r. \quad (6)$$

It seems quite reasonable that the ratio v/c should occur explicitly.

(ii) When we use Gaussian definitions, analogous electric and magnetic quantities have the same dimensions (even in a four-dimensional system); a magnetic pole has the same dimensions as an electric charge; H has the same dimensions as E . (This is not true with the usual definitions.) Moreover, we find analogous formulas. The force between two poles in vacuum is $m_1 m_2 / kr^2$, analogous to $q_1 q_2 / kr^2$ for charges, where the constant k depends on the units used; it is unity in Gaussian units.

(iii) There is a real distinction between "electric quantities" and "magnetic quantities" which is generally recognized. The basis of this distinction is hard to find when non-Gaussian definitions are used. Magnetomotive force measured in ampere turns and magnetic flux measured in volt seconds per turn have nothing to distinguish them from electrostatic quantities. The quantity c occurs in all Gaussian magnetic quantities as a distinguishing symbol, indicative of their relativistic nature.

The Gaussian magnetic equations can be simplified in formal appearance if we adopt the symbol i for i/c —that is for $\eta v/c$, calling i the *reduced current*, or "the current in electromagnetic units." This i is dimensionally not a current but a linear charge density. The force between

two parallel currents i_1, i_2 can be expressed as

$$F = \frac{2i_1 i_2}{r} l;$$

comparison with Eq. (6) shows that i is the magnetic analog of q . The reduced current may be regarded as the fundamental magnetic quantity. The definitions in Table I are given in terms of i instead of i and c explicitly.

There is no necessary connection between this ambiguity of definition for magnetic concepts and the choice of fundamental dimensions. Indeed, some ostensible "three-dimensionists" adopt the non-Gaussian definitions and refer to an "electrostatic system of magnetic units" and to a "magnetic system of electric units." However, these systems are seldom used. The fundamental reason for the three-dimensional system is simplicity; the simplest unit of charge is the statcoulomb, and the simplest unit of magnetic field is the gauss. Hence the only complete system of three-dimensional units in general use is the Gauss system, using the Gaussian definitions. Thus in actual practice the competing systems are the Gaussian three-dimensional system generally used in physics and a four-dimensional system (such as the mks) with non-Gaussian definitions. Physical equations in the two systems differ by a factor k and, in magnetic equations, also by a factor c .

The Gaussian system requires no discussion among physicists. We turn to the four-dimensional systems (mks, cgs, cgsa).

Four-Dimensional Systems

Charge is to be considered as a fundamental unit. Coulomb's law for empty space is then

$$F = q_1 q_2 / k'' r^2,$$

where k'' is a conversion factor with a value depending only on the system of units used. It is equal to 1 statcoul² sec² g⁻¹ cm⁻³, or 1/(9×10²⁰) abcoul² sec² g⁻¹ cm⁻³, or 1/(9×10⁹) coul² sec² kg⁻¹ m⁻³.

The constant k'' , called "the dielectric constant of space," is dimensionally [Q²T²/ML³]. If, by a change to a smaller charge unit, q is increased n -fold, k'' will be increased n^2 -fold.

We shall say that k'' is doubly covariant, referring to its variance with respect to q .

It will be important to distinguish between quantities that are defined as proportional to q (covariant quantities) and those defined as proportional to $1/q$ (contravariant quantities). The former we shall designate by a dot superscript and the latter by a dot subscript. When the charge unit is increased by a factor n , the unit of a covariant quantity is increased by a factor n and the unit of a contravariant quantity is reduced by $1/n$. In any equation covariant and contravariant quantities must balance. For example, of the equations $V = q'/r$ and $V = q'/k''r$, each of which is numerically correct in cgs units, only the second is dimensionally correct.

Any quantity can be defined covariantly or contravariantly (always referring to its variance with respect to q), but to avoid ambiguity quantities are actually defined only in one sense. Current is naturally defined [(c), Table I] covariantly by the equation $i = q'/t$, the practical unit being 1 coul sec⁻¹. Potential [(d), Table I] in empty space might be defined as $V = q'/r$ (due to a point charge) or as $V = W/q'$ (work per unit charge). The covariant definition describes the potential field in terms of its origin; we may call this the "generated potential." The contravariant definition describes it in terms of what it does. This is the "acting potential." The latter contravariant definition is the one which is used, and the practical unit of potential is the joule per coulomb (not the coulomb per meter).

When charges are immersed in a fluid the force between them is reduced [(b), Table I]; the ratio of reduction determines the relative dielectric constant ϵ . More generally (for solid dielectrics) ϵ can be determined by noting the increase in capacitance of a condenser in which the dielectric is substituted for vacuum. The relative dielectric constant is dimensionless. The product $k''\epsilon$ is called the absolute dielectric constant.

There are four field concepts that must be recognized, though they are not all formally named. In addition to the dimensional difference which we indicate by dots, there is a structural difference which we indicate by letters (E, D). The E -field is irrotational under electrostatic conditions, and the D -field is sourceless in the

medium (outside of charges). We shall call them "the irrotational field" and "the sourceless field," referring to the structure only of electrostatic fields in the medium. The ratio of the two fields (measured in the same units) is the relative dielectric constant ϵ .

It would be possible to define either E or D covariantly or contravariantly. The definitions⁵ are given in Table I, (e), (f). A little consideration will show that the quantity q'/ϵ which appears is no less fundamental than q' . There are electric poles m'_e (which we can also call "apparent charges") on the surface of a polarized dielectric. These will be defined later. When a charge is surrounded by the medium, the induced pole is $q'(1-\epsilon)/\epsilon$. If $Q' = q' + m'_e = q'/\epsilon$, including "bound surface charges" with free charges, then the field due to a point charge is

$$D' = q'/r^2 \text{ and } "E'" = Q'/r^2. \text{ (Generated fields.)}$$

The contravariant definitions are given in terms of the force on a charged body:⁶

$$"D." = F/Q' \text{ and } E. = F/q'. \text{ (Acting fields.)}$$

Actually the irrotational field (E) is always defined contravariantly, as an acting field ($E.$); the sourceless field (D) is always defined covariantly, as a generated field (D'). There are no specific symbols for our E' or $D.$; these are represented by $k''E.$ and D'/k'' , respectively. Or alternatively, they may be represented by D'/ϵ and $\epsilon E'$. As an example, the equation commonly written $D = E + 4\pi P_e$ is dimensionally incorrect;

it must be written $D' = k''E. + 4\pi P'_e$. The relations among the four fields are shown in Fig. 1; here ϵ''_e is the "absolute dielectric constant" $k''\epsilon$.

Turning to magnetism, we have now, instead of Coulomb's law, the law for the force between parallel currents in a vacuum:

$$F = \frac{2i_1 i_2}{k''c^2 r} l = \frac{K.. 2i_1 i_2}{r} l.$$

Here $K.. = 1/k''c^2$; the factor c^2 is included in the constant since we do not wish it to appear explicitly. This constant $K..$ is called "the permeability of empty space." It is evidently doubly contravariant in i' and hence in q' ; it is numerically $1/(9 \times 10^{20})$ in cgss units, 1 in cgss units and 10^{-11} in mksc units.

If the two wires are in an infinite homogeneous fluid the force between them is changed by a factor μ , the relative permeability. More generally (for solid mediums), μ can be measured by noting the increase in mutual inductance. This factor μ is invariant; $K..\mu$ is called the absolute permeability, $\mu_a...$

The first quantities in magnetism to be defined, after this universal constant $K..$ and this dimensionless constant μ , are the magnetic fields. Again there are four concepts, though only two of them are dignified by symbols. In addition to the dimensional difference there is a structural difference between a B -field, which is sourceless, and an H -field, which is irrotational (in the medium, outside of currents). We consider first the fields in empty space (where B and H are not structurally distinguishable). In this case, H' (or " B'' ") $= 2i'/r$, for the field about a straight wire; $F = i' l B$. (or " H ." for the force on a wire. In mksc units the covariant (generated) field is expressed in amperes per meter; the contravariant (acting) field is expressed in webers per square meter. The corresponding cgss units are the oersted and the gauss. In space, the number of oersteds is equal to the number of gauss.

The elementary magnetic dipoles are circular currents, and in a magnetized body these give a resultant "apparent current" j' around the periphery. If a wire is in a magnetic medium the apparent current j' at the surface is equal to $i'(\mu-1)/\mu$. If we let $I' = i' + j'$, including currents and apparent currents, we have for the

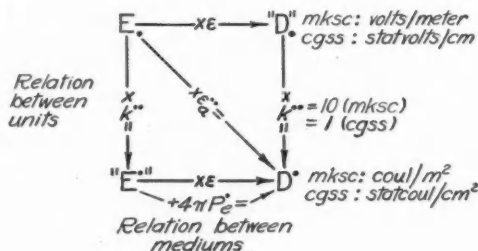


FIG. 1.

⁵ The definitions in Table I are chosen for simplicity rather than generality. For example, (e) and (f) do not apply to a nonhomogeneous medium. The defining example used implies that $\text{div } D = 4\pi$ (charge density); the special case does not suggest that $\text{curl } D = 4\pi \text{ curl } P_e$.

⁶ The first equation is more general than the second; it also gives the force on the pole of an electret.

magnitude of the field about a long straight wire,

$$"B" = 2I/r \text{ and } H = 2i/r. \text{ (Generated fields.)}$$

The contravariant definitions are given in terms of the force on a straight wire,

$$B = F/i l \text{ and } "H" = F/I l. \text{ (Acting fields.)}$$

The force equation is frequently written $F = i l H$ (instead of $i l B$); this is wrong even in a three-dimensional system, but the error is then trivial since μ is practically unity in any fluid.

In actual practice only the H -field and the B -field are formally recognized. The other two fields, when required, are represented as follows: " H ." as $H K..$ or $B./\mu$ (gausses); " B ." as $B./K..$ or μH (oersteds).

The familiar equation which, in three-dimensional notation, is written

$$B = H + 4\pi P_m,$$

is now dimensionally incorrect; it must be written

$$B = K..H' + 4\pi K..P',$$

if we assume that the intensity of magnetization P_m is defined covariantly. It is evident now that the equation $F = i l H$ is wrong not only because it introduces the field of wrong structure but (more seriously) because the field has the wrong dimensions and gives entirely the wrong answer except in cgsa units.

By multiplying H' by the absolute permeability $\mu_{a..} [= K..\mu]$, we can change both units and structure in a single step and so derive B , immediately from H' . For example, suppose we have a long iron rod of absolute permeability 10^{-8} mksc units (relative permeability 1000) in a long cylinder around which flows a current with linear density of $\frac{1}{4}\pi$ amp m^{-1} . The H -field in the rod is 1 amp m^{-1} ("practical oersted"), and the B -field is 10^{-8} weber m^{-2} . This obscures the actual role of the iron. In the iron, the field strength (H') generated by the cylinder current is 1 amp/ m^{-1} ; around the iron rod surface then flows an apparent current 999 times the cylinder current, and the field generated by this is 999 amp m^{-1} , giving a total generated field (B') of 1000 amp m^{-1} . The factor $K..$, which is now introduced to change the units to webers per

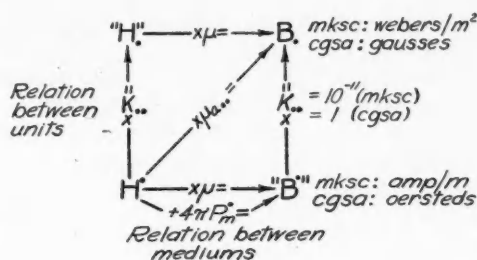


FIG. 2.

square meter, has nothing to do with the magnetization of iron.

The magnetic moment of a solenoid can be defined by the equation $M = NiA$; or by the expression for the torque on it when it is perpendicular to a magnetic field, $T = M \cdot B..$ These are both covariant definitions. The possible contravariant definition, $T = M \cdot H' \epsilon$, is obviously quite artificial. Magnetic moment of a solenoid must be regarded as a covariant quantity; its practical units are ampere turns per square meter.

Except for the dielectric constant and the permeability, we have not defined the properties of materials. As regards electric properties it will suffice to remark that they are defined as covariant in the charge. In regard to the magnetic quantities—magnetic moment, intensity of magnetization and pole strength—there is a variation in usage. Intensity of magnetization P_m is defined as the magnetic moment per unit volume, and pole density is defined as $-\text{div} P_m$. Hence M , P_m and the pole strength m all have the same variance. It would seem reasonable to define the magnetic moment of a bar magnet as equal to that of a solenoid which produces the same external field and experiences the same torque in a vacuum. Its magnetic moment would then be " I " A , where " I " is the apparent current around its perimeter (the integrated effect of the dipole currents), and A is its cross-sectional area. These quantities would then all be covariant. A typical equation—that for the force on a pole—then would be

$$F = m \cdot B, \quad (7)$$

or possibly $F = m \cdot H \cdot K..$, depending upon which

was actually true in a medium. This is the usage of Stratton.⁷

However, this is not the general convention. These properties of materials are usually regarded as contravariant (defined in reciprocal amperes), and the equation for the force on a pole is given as

$$F = m \cdot H, \quad (8)$$

and the field of a point pole as

$$B = m/r^2. \quad (9)$$

Thus the H -field (for example, in oersteds) appears here as an "acting-on-magnets" field, and the B -field (gausses) as a "generated-by-magnets" field, though the H -field (oersteds) is a "generated-by-currents" field and the B -field (gausses) is an "acting-on-currents" field (illustrated by such equations as $H = 4\pi ni/l$, $F = i\ell B$, $\mathcal{E} = d(BA)/dt$, and so forth). This is likely to produce confusion. For example, the subject of terrestrial magnetism is at best rather complicated when we introduce two fields—a gauss field and an oersted field—of the same numerical value. But if a teacher adopts the usual convention, in presenting the subject consistently he must picture the original, unreduced field as possibly an H -field (if due to terrestrial material) which must be reduced to gaussses to find the force on a current or as possibly a B -field (if due to earth currents) which must be converted into oersteds to find the force on a pole, or as actually something of both, which must be converted both ways for the force on an electromagnet. It is a little confusing.

The reason for the contravariant definition of m , which we have seen to be artificial, was the presumption that Eq. (7) was in error and that Eqs. (8) and (9) were correct. It was supposed that the force on a permanent magnet in a medium depended on the irrotational rather than the sourceless field. If the pole is defined contravariantly, the dimensional constant K does not appear in Eq. (8). The consideration is trivial because B is always practically equal to H in all fluids. However, it has been shown that Eqs. (8) and (9) are not in general correct.⁸

Moreover the writer has shown that Eq. (7) is correct and also that $B = m/r^2$ for the exterior field of a permanent magnet.⁹ Hence there seems to be no reason for the artificial convention. The usage of Stratton should be generally adopted. The equation for the force on a pole is $F = m \cdot B$, and that for the force on a current is $F = i\ell B$.

There is considerable confusion about the use of four-dimensional electrical units; this pertains particularly to the field concepts. The difficulty is that we have inherited a symbolism in which "fields of force" were represented by H instead of B ; we have retained it since the error does not at once betray itself by a wrong result when we use cgsa units, as is customary in physics. The units have followed the letter, and it has become fashionable to express magnetic fields quite generally in oersteds. In a large majority of instances the field being considered is a field of force and the unit is the gauss.

To avoid this confusion we should (i) distinguish between covariant and contravariant concepts, (ii) define poles covariantly, (iii) check equations for a balance between covariant and contravariant quantities. Even so a student is likely to become confused. This confusion will be reduced if we emphasize *one* field, the acting field, or "field of force." There is no confusion in gravitational theory, since the gravitational field is defined only as the acting field; $g = Gm/r^2$, not m/r^2 . In electricity the D -field is ignored in elementary teaching. Possibly in ferromagnetism the H -field cannot be ignored entirely, but it can be regarded as an intermediate step on the way to finding the B -field. Only in the theory of radiation are the B - and H -fields (and the E - and D -fields) treated as of equal importance.

The student is often introduced to electricity and magnetism through study of the magnetism of the earth. Certainly we need here refer only to

⁸ This is unpublished. Briefly the argument is as follows. It is a matter of choice whether "apparent charges" m , and "apparent currents" be included in the concepts of charge q and current i ; by definition they are excluded, charge density being $\text{div} D$, not $\text{div} E$. But we must regard the pole on the magnet (or electret) and that in the medium as the same concept; the total pole strength, $\int \text{div} P d\tau$, includes both. With this recognized, it is found that the roles played by B and H (or D and E) in the familiar equations are exactly reversed. The equations obtained ($F = mB$ and $H = m/r^2$ for point poles) are shown to be generally valid, in contrast to the usual equations.

⁷ Stratton, *Electromagnetic theory* (McGraw-Hill, 1941).

⁸ L. R. Wilberforce, *Proc. Phys. Soc. London* **45**, 82 (1933); **46**, 312 (1934); L. Page, *Physical Rev.* **44**, 112 (1933); Page and Adams, *Am. J. Physics* **3**, 51 (1935); C. C. Murdock, *Am. J. Physics* **12**, 201 (1944).

the B -field, measured in gauss, which acts on our compass needle. This is the final field produced by the earth. There may be an intermediate H -field used in deriving it, an oersted field, an uncompleted field which requires the introduction of K to fit it for use; but it only creates confusion to introduce it. In Fleming's

Terrestrial Magnetism and Electricity (1939) the oersted is not mentioned. In Chapman and Bartel's *Geomagnetism* (1940) it is mentioned only to be ignored. It can be said that a teacher who introduces an oersted field into the discussion of terrestrial magnetism violates the most authoritative convention.

The Logic of Quanta

GUSTAV BERGMANN

Department of Philosophy, The State University of Iowa, Iowa City, Iowa

"THE philosophical problems of quantum mechanics are centered around two main issues. The first concerns the transition from causal laws to probability laws; the second concerns the interpretation of unobserved objects." All philosophical analysts will agree with this statement of Reichenbach's in the opening paragraph of his recent book¹ as to what they hope to accomplish when they turn their attention to the momentous development of quantum mechanics that took place during the last two decades. But unanimity about the *what* does not imply unanimity as to the *how*. I, for one, am unable to accept the fundamental ideas of Reichenbach's approach; consequently, I find myself in disagreement with several important features of his analysis. Perhaps the present essay would not have been written, were it not for the challenge of his work to all those who, without sharing his philosophical ideas, share his conviction that the philosophy of science, if it is to be significant, must be minute, technical, and—in a loose sense of the terms—operational and positivistic. Reichenbach's book is all this and, to that extent, very admirable. Yet I shall here present my own analysis quite independently and, but for a few incidental remarks, quite unpolemically. I wish to say, however, that Reichenbach's work is one of the two from which I learned most. The other is von Neumann's *Mathematische Grundlagen der Quantenmechanik*.

¹ H. Reichenbach, *Philosophical foundations of quantum mechanics* (Univ. of California Press, 1944).

When I speak of the logic or philosophy of science, or of a certain area of science, I use the terms as I did in some previous papers that appeared in this journal.^{2,3} Hence I shall not take time to explain them again in any detail. Briefly, the logical analysis of science is a certain kind of description of science. It is not technical about the things science is technical about, yet it has standards and a technic of its own. To give an illustration, a logician who uses such words as 'causal,' 'operational,' 'measurement,' must try to be as precise as physicists are in speaking about force and energy. The purpose of the present essay is to outline a description of the modern quantum theory that is technical and, I hope, precise in this sense. The job is still largely undone. Reichenbach's book is the first systematic attempt that goes beyond the obvious generalities.⁴ Physicists have given us many valuable hints; but these are mostly isolated flashes of insight. Some other comments they have made are either vague or outright confusing. If, for instance, I read, in a physicist's book on quantum mechanics, that it is "operationally meaningless" to speak of the particle's position

² G. Bergmann, "Outline of an empiricist philosophy of physics," *Am. J. Physics* 11, 248, 335 (1943).

³ G. Bergmann, "The logic of probability," *Am. J. Physics* 9, 263 (1941).

⁴ This is not meant as a criticism of Philipp Frank's *Interpretations and misinterpretations of modern physics* (Hermann & Cie., 1938), a painstaking and lucid refutation of the various attempts to build a metaphysics of freedom upon the vocabulary of modern physics. But the significance and intention of this book are cultural rather than technical.

and momentum beyond the limits of the indeterminacy relation, I see immediately that the term is not used in any precise and specific sense. So I react as a physicist would react when a speculative philosopher tries to buttress his constructions by an analogical use of 'force' and 'energy.' So-called operational analysis is but one aspect of the description of science and applies only to *empirical constructs*. Appeal to it is quite out of place when one examines those features of a calculus or *computational schema* that determine whether it is a *model*, that is, the kind of schema in which one can—in a certain sense—speak of the "position" and the "momentum" of a "particle." Again, a model's being meaningful does not depend on the possibility of assigning precise numerical values, or any numerical values, to all its magnitudes by means of the *theory* of which it is a part. At any rate, such a criterion of meaning would throw out theories which, though now obsolete, are methodologically quite acceptable.

I wish that I could now define the four italicized terms—empirical construct, computational schema, model, theory—and show the advantages such definitions yield, by way of clarification and through the elimination of pseudo-problems. Unfortunately, a moderately adequate explanation fills 30 columns of this journal.² Nor can I, with respect to probability, repeat the exposition and defense of the frequency view that I have undertaken elsewhere. But the reader will, in the first part of the present paper, find as much of this material as could be included without untoward repetition. The first part (SECS. I to III) deals with the fundamental notions of measurement, determinism and mechanism, causality, probability, and statistics. Except for a scattering of anticipatory remarks to hold and direct the attention, these sections contain no direct references to the quantum theory. Taking advantage of the background thus provided, the second part (SECS. IV to VII) begins with an outline of the general structure of the quantum theory. This is followed by a description of the calculus whose construction was begun by Schrödinger and Heisenberg, and an analysis of the position that this calculus—the quantum-mechanical schema—occupies in the total structure of the theory. The last section is a separate discussion of two important problems in the logic of quanta.

I. Measurement

(1) Speaking within physics rather than philosophizing about it, we use the term 'measurement' very broadly. We say that we measure the temperature of a gas, but we also say that we measure the (average) velocity of its molecules. These are two different things. The difference I have in mind is not that in the first case we simply read an instrument, while in the second we derive the numerical value from several such readings through a fair amount of computation. The important difference is, rather, that in the case of temperature we measure an empirical construct, while the second number⁵ receives its full meaning or interpretation only as the result of an additional step, the coordination of, say, the classical kinetic model to the empirical constructs and laws of thermodynamics. Ambiguities of this kind are harmless so long as they do not obstruct the precise description of situations where the difference they neglect does make a difference. With respect to 'measure' and 'measurement,' such a situation has arisen in the quantum theory. If, therefore, I shall hereafter speak of 'measurement,' the expression will refer to the *measurement of empirical constructs*; whenever the magnitudes of a model or certain other schemas are involved, I shall speak of the *assignment of numerical values*.

(2) In some discussions of the quantum theory one reads that observation and measurement, by affecting the objects observed or measured, introduce an ineradicably *subjective* factor into all our knowledge. Having delivered himself of this opinion, the writer usually pays a compliment to the positivists or other philosophers whom he believes to have disposed of the fable of an independent or objective reality. All this is very confusing. What the physicist impugns, or has a right to impugn, is not the common-sense realism that we all share but merely the naive belief in the "reality" or "existence" of the classical models. The particles of these models are not, and

⁵ Since it is computationally derived from the values of empirical constructs, this number is, in one sense, also an empirical construct. But this sense is trivial; it neglects differences that we feel exist and that should be clarified. An *empirical construct* is—very roughly speaking—a concept defined in terms of the immediately observable by a chain of "explicit definitions." See also SEC. II (4), where it is contrasted with the "theoretical constructs" of a model.

never were, either observed or unobserved objects, in the only intelligible sense of the term—the sense in which trees, stones, laboratory equipment are objects that may or may not be observed. The confusion stems thus from two sources: first, from an analogical use of ‘observed’ in such phrases as ‘indirectly observed’, second, from the unguarded use of ‘subjective’, ‘objective’, ‘reality’ and similarly slippery terms that are better left to the philosophers. (Philosophers of my stripe try very hard to avoid them.)

Measurement is based on the observation of scales, and I have never heard it suggested that we make a needle move by watching it, which is but another way of saying that on the common-sense level of laboratory objects and their immediately observable properties and relations the language of common-sense realism is the only reasonable one. The ancient puzzle how we can ever know that when we do not watch them stones, trees, and laboratory equipment behave exactly as when we do watch them is better left to the epistemologist. This is a *verbal* puzzle, not a *facutal* assumption that we may be induced to drop by developments, no matter how revolutionary, in theoretical physics. Conversely, the solution of this puzzle—if it is a puzzle—depends on the *grammar* of ‘perceive (observe)’ and ‘exist,’ not on the *physiology* of perception, with or without benefit of the indeterminacy relation. To drag physiology or the so-called mind-body problem into the methodological analysis of physics makes no more sense to me than Descartes’ “derivation” of the conservation of momentum from the perfection of the Creator. To say the same thing again, in three explicit steps: (i) our observing them makes, on the common-sense level, no difference to the behavior of ordinary physical objects; (ii) this behavior is the subject matter of all measurement and, in a familiar interpretation, of all science; (iii) therefore, it is hard to see how a philosophical or otherwise radical notion of subjectivity could enter into the methodological analysis of science. These are the reasons why I believe that whenever this particular twist has been introduced, it has, far from contributing to it, retarded a precise description of the quantum theory. To quote Reichenbach who, for reasons of his own, makes the same point: “Like all other parts of physics,

quantum mechanics deals with nothing but relations between physical things; all its statements can be made without reference to an observer.”⁶ But then, for Reichenbach a tree and an electron are both “things.”

Confused as the talk about subjectivity is, it often appears in remarks about a very fundamental aspect of measurement. Systematically, this aspect is not quite new nor, as a point about measurement, peculiar to the quantum theory. Historically, it seems that this aspect did not receive all the attention it deserved before it was brought up—though often for the wrong reasons and in the wrong context—in recent discussions of the quantum theory. I turn now to the discussion of this aspect.

(3) In measuring an empirical construct exemplified by an object or situation *A* at a given moment—or, as I shall say briefly, in measuring *A*—one does not observe *A* alone but, rather, certain aspects of a situation (*A, B*), compounded of *A* and the yardstick or measuring instrument *B*. There is thus the possibility of an interaction by which the two components of the new situation, *A* and *B*, may produce changes in each other. That gives rise to two questions: (i) how can we recognize such changes? (ii) under what conditions is a feature of (*A, B*) acceptable as a measurement of *A*, that is, as an index or characterizer of *A* alone?

The answer to the first question is self-evident. We shall say that *A* has been changed by being put in the measuring situation if it subsequently behaves *in some respect* differently from *A'*—which is *otherwise exactly like A*, but has not been measured—provided that *the difference cannot be attributed to other factors*. If differences occur only while (*A, B*) is maintained, the change may be called temporary. The three italicized clauses reveal the hypothetical and conventional ingredients that inhere in the notion of change. Their precise discussion, though very laborious, is sufficiently elementary and familiar that we need not bother with it in this paper. The point is simply that whatever we know, on the levels of either science or common sense, we know not by itself but embedded, as it were, in a network of hypotheses (empirical laws) and defi-

⁶ Reference 1, p. 15.

nitions. If a measurement does not, in this sense, produce a change in A , then its immediate repetition will yield the same value. Clearly, the converse does not hold; there is still the possibility of different or later changes. In accepting the repetition test as a *criterion of noninteraction*, one introduces, therefore, but another convention—though, of course, a plausible one. However, the repetition test has been suggested, not as a criterion of noninteraction, but as part of the *definition of measurement*. The suggestion is unnecessarily restrictive. Perhaps it would not have been made if the distinction between measurement proper and the assignment of numerical values to the magnitudes of the quantum-mechanical schema had been appreciated. But of this I shall speak later. Let me now try to give a nonrestrictive definition of measurement. To do this is to answer the second of the two questions I have proposed.

A property of (A, B) is a measure of A if and only if it enters, together with other such properties of A (and of other objects), into empirical laws that predict or postdict the behavior, before or after the occurrence of (A, B), of A (or of A in interaction with other objects). To be a measurement in this sense, a property of (A, B) need not satisfy the repetition criterion; yet it seems to me that the definition realizes the scientist's working conception of measurement. To show this, assume that there is, in a world otherwise like ours, a magic rod, such that any expanse laid off against it contracts, virtually instantaneously and permanently, to nine-tenths of its previous length. Since this world is by assumption otherwise like ours, its physicists could discover the curious phenomenon. They would know that an object which, thus measured, is found to be of length l will henceforward behave like such an object; and they would also know that up to the time of the measurement with the magic rod it behaved like an object of length $10/9l$. I wonder what else one could expect measurement to achieve. As a matter of convenience, measurement without interaction is often preferable. But such measurement is not always available. One merely has to think of the measurements of temperature and of electric field strength by means of thermometers and electroscopic devices. Finally, whether or not there is interaction does not depend on the

schema, classical or otherwise, that we use in our theories. Measurement, as here defined, is entirely a matter of empirical constructs and empirical laws.

Presently I shall define what we mean by 'accuracy' and 'statistical.' In order not to have to return to the discussion of interaction, I shall first make two remarks that are probably clear already; they will, at any rate, be clear after these further notions have been introduced. (i) Interaction in measurement does not, as such, impose any limits on its accuracy. (ii) Interaction in measurement does not prejudge the form of the empirical laws, whether statistical or nonstatistical, in which the constructs measured occur; nor does it, *a fortiori*, determine whether the theoretical schemas that account for these laws are either statistical or nonstatistical. If the electron and the photon of the classical Compton experiment were ordinary physical objects such as stones and light beams, one could therefore say, as Reichenbach does, that "the disturbance by observation, in itself, does not lead to the indeterminacy of the observation."⁷ However, by thus neglecting an important difference this formulation blurs another distinction. The indeterminacy relation, which is a matter of the *quantum-mechanical schema* and not of observation, is *one* thing. The impossibility of assigning, within the *semiclassical model*, precise values to both the position and the momentum of the Compton electron before the collision is *another* thing. Again, this impossibility has nothing to do with interaction in, or the limits of accuracy of, measurement. It is, as we shall see, a consequence of certain empirical laws and of the manner in which the entities of the semiclassical theory are coordinated to the empirical constructs. By 'semiclassical theory' I mean the schema—devised by Bohr, Einstein and others—that was in use before Heisenberg and Schrödinger introduced their radical innovations.

(4) We are, as Bertrand Russell once said, quite certain that Cleopatra had 2 eyes and 1 nose and not, perhaps, 2.000001 eyes and 0.999998 noses. Counting is, indeed, the only empirical use of numbers that is precise and accurate. The measurement of continuous dimensions,

⁷ Reference 1, p. 17.

such as time and space, is not precise and accurate in the same sense. Familiar as this sounds, it is, I believe, worth while to state carefully what we really mean by such terms as 'precise' and 'accurate.' Before turning to this task, I wish to make two preliminary remarks. They, too, cover familiar ground. *First*, the use of real numbers and continuous functions in the description of the empirical material is but a very successful convention, adopted because of the many conveniences it affords in formulation and computation. Accordingly, 'continuous' as just used in 'continuous function' is a purely arithmetic notion, to be distinguished from the 'continuous' in 'continuous dimension,' a phrase that occurs above. In the latter usage 'continuous' is a qualitative term that does not require further analysis for our purposes. *Second*, there is no direct connection between the 'precision' of the discrete eigenvalues that occur in the quantum-mechanical schema and the 'precision,' or the lack of such, with which we measure position, breadth and intensity of the spectral lines whose behavior we explain by a theory of which that schema is a part. It is, of course, a consequence of this theory as a whole that there are limits to the accuracy with which we measure empirical constructs. But then, such limitation is also a consequence of the semiclassical or the classical theory or, for that matter, of almost any corpuscular and, in this sense, discontinuous theory of matter. Obvious as all this is, it cuts the ground from under a good deal of loose talk.

One may measure the length of an iron rod with an ordinary yardstick to the nearest full inch, or one may measure the same stick with a more elaborate instrument to the nearest 0.01 in. In either case, as in all measurement, one manipulates physical objects and, eventually, reads a scale. The perceptual exertion required may actually be greater in the first case than in the second. Yet we call the second measurement more precise than the former—or this, at least, is how I shall define 'precision.' *Precision*, then, means *the number of digits of a given unit*. The larger this number, the greater the precision. How precise we can be is a matter of empirical laws and, in particular, of those empirical laws that are sometimes referred to as the theory of the instrument. On the other hand, a measure-

ment whose precision is much less than the best we can do may be completely reliable, a measurement being called *reliable* when *in a large number of repetitions the result is always the same*. (The qualifications due to possible interaction and other factors that may interfere with this sameness need not bother us after what has been said before.) If the necessary care is taken, the first of the two measurements of the iron rod is, in fact, completely reliable. The second measurement, which is more precise, is less likely to be completely reliable. The values obtained will scatter or, as one also says, their standard error will not be equal to zero. Having thus defined precision and reliability, I turn to a definition of *accuracy*. The following is, I believe, an exact statement of that rather fundamental feature of our world to which we refer when we say that there is, in fact, a limit to the accuracy of our measurements. *A measurement as precise as we can make it is never completely reliable. Its standard error, though absolutely decreasing with increasing precision, shows no tendency to decrease in proportion to the last digit*. Conversely, if our most precise measurements were completely reliable, we would not consider them as of limited accuracy. Such a state of affairs, by the way, would not necessarily imply that we must forgo the convenience of real numbers in describing it. Whether at all and under what conditions it would be compatible with a corpuscular theory is not an easy question; it has, I believe, some connection with certain aspects of the Zeno paradoxes.

As is well known, we do not in careful experimental work expect our measurements to be reliable. We repeat them, define their average as the "true value" and operate in the formulation and testing of laws with the value thus obtained. Anybody who wishes to describe this state of affairs by saying that all laws of nature are "statistical" is free to do so. But having made this choice of meaning, he is no longer free to use the same term in a different and more specific sense in which not all but only some empirical laws and theories are statistical. Or, at least, he may not do so without being explicit about it. Furthermore, anybody who is thus explicit will not be tempted to believe that the inaccuracy of measurement, by making all laws "statistical," implies or even suggests the "statistical" nature

of the quantum theory. In order to make this point as vigorously as possible, I shall henceforth assume that our measurements can be made ideally accurate. The assumption is, of course, contrary to fact. Its expository value lies in the circumstance that with it the methodological problems of the quantum theory are exactly what they are without it.

II. Mechanism and Determinism

I turn now to questions that are usually considered under the headings "determinism" and "mechanism." It will be best to begin with an illustration. Assume, then, that a physicist observes, over a period of time, a *configuration* consisting of n objects, P_1, P_2, \dots, P_n . From previous observations he knows that each P_i is characterized by a number m_i , the value of an empirical construct called the mass of P_i , so that $dm_i/dt = 0$. The set of numbers $[m_1, m_2, \dots, m_n]$ is a construct that one may use to characterize the configuration. At a certain moment ($t=0$) the physicist measures certain empirical constructs characteristic of the *condition* of the configuration at that moment, namely, the $3n$ position coordinates (q_1, q_2, \dots, q_{3n}) and the $3n$ momentum coordinates (p_1, p_2, \dots, p_{3n}). The $7n$ numbers m_i, q_i, p_i , when substituted in a certain formula or computation rule, yield $6n$ functions of a continuous parameter, $q_1(t), \dots, p_{3n}(t)$. (The computation rule prescribes, of course, the solution of the canonic equations: $\dot{p}_i = -\partial H/\partial q_i, \dot{q}_i = \partial H/\partial p_i$.) The $6n$ values q_1^1, \dots, p_{3n}^1 of these functions for $t=t_1$ are then the values the physicist predicts ($t_1 > 0$) or postdicts ($t_1 < 0$) for the coordinates of P_1, P_2, \dots, P_n at the time t_1 . This is, of course, a philosopher's description—with its characteristic shift in emphasis—of the classical treatment of the problem of n bodies. I shall use it to introduce some distinctions.

Let us first consider the whole thing formally, as a definitional arithmetic *schema* of the sort that is also called a calculus. First, one defines three kinds of entities: systems, states, processes. A *system* is an ordered set of M real numbers $[c_1, c_2, \dots, c_M]$. A *state* is an ordered set of N real numbers $[x_1^0, x_2^0, \dots, x_N^0]$; and there is also a rule that determines when a state may be said to be a (possible) *state of a system*. In our illustration, systems are restricted to sets of positive

numbers and the rule is, very simply, $N=6M$. A *process* is an ordered set of N functions—with certain mathematical characteristics—of a continuous parameter t , which is referred to as time, $[x_1(t), \dots, x_N(t)]$. Next, given a system and one of "its" states $[x_1^0, x_2^0, \dots, x_N^0]$, there is a rule that allows for the computation of one and only one process so that $x_i^0 = x_i(0)$. This rule is the part of the whole schema that is usually spoken of as the calculus; I shall refer to it as the *process formula*. It has, typically, the form of a system of ordinary differential equations. Of the schema as a whole I shall speak as a *process schema*. I could also call it a "causal" schema, but I prefer not to make any technical use of this term.

Several generalizations and variations of the process schema are of interest. In one rather common case, states and systems are not defined as finite ordered sets of numbers, but as functions, or ordered sets of functions, with certain mathematical characteristics, of one or several continuous variables. Such process schemas are called *field schemas*. Their formula has, typically, the form of a system of partial differential equations. Again, if these equations are of a certain special form, the schema is called a *wave schema*. These are the only precise meanings the terms 'field' and 'wave' have in theoretical physics.⁸ To mention another possibility, the process formula may allow for the computation of a process not from one of its states, but only from one of its segments ($a \leq t \leq b$). It is then, typically, a system of integro-differential equations. As Volterra first pointed out, such a schema represents our notion of "historical" lawfulness in one precise sense of the term. In another, weaker sense, all process schemas are historical.

It will be noticed that I have distinguished between configurations and conditions "out there," on the one hand, and systems and states, which are purely arithmetic entities, on the other. Also, I have spoken of schemas (calculi) and, in doing so, avoided the customary term, theory. There is a point to all this circumstantiality. For instance, a certain type of theory⁹—I shall call it *process theory*—can now be described as follows. (i) One designates a class of configurations. For a theory

⁸ Some confusion has recently been created in the methodology of the behavior sciences (psychology, sociology) through an ambiguous use of the term 'field.' While some analogies with physical field theories are claimed, the term actually is used in the sense of interaction.

⁹ In my earlier paper (reference 2) the term 'theory' was restricted to what is here called a theory with a mechanical model; other "theories" were spoken of as systems of empirical laws. The present change in terminology makes it easier to use 'system' in the way in which it is customarily used in quantum mechanics.

to be considered worth while, this class must be rather comprehensive. Let C be an arbitrary member of it. (ii) One defines a finite class E of kinds of empirical constructs realized in these configurations. (iii) One selects a set of measurements of constructs of E to be performed on C and on any of its momentary conditions. Let the numbers of these measurements be M and N , respectively. (iv) One defines a process schema S . (v) One establishes a special set of rules—the so-called coordinating definitions—such that from each set of $M+N$ measurements, the last N of which characterize the condition of C at time t_0 , one and only one set of N numbers can be computed, by means of S , for an arbitrary value t of the parameter. If this latter set, so far as we know, is always identical with the measures that characterize, in proper order, the condition of C at time t , then the theory is said to be successful.

If the theory is successful, then two configurations that yield, literally, identical measurements—this identity is one of numbers—are said to be identical; similarly for conditions of identical configurations. *Identity* in this sense is clearly a matter of definition. It signifies what I expressed, in SEC. I, by speaking of two objects or situations as exactly alike. Also, it is merely an identity with respect to a given theory and cannot even be stated without reference to the success of the latter. The behavior of configurations that allow for such treatment may be said to be “determined” by any of their temporal cross sections (initial conditions). Indeed, we are now ready to select a precise meaning for the term ‘determinism.’ But, of course, any such choice of meaning is arbitrary in that rather limited and peculiar sense in which one cannot help being arbitrary when one tries to state precisely and abstractly what people speak and think of in terms of concrete cases. For instance, ‘mechanism’ and ‘causality’ are sometimes used to connote, among other things, what is here called determinism. What words are used is of no consequence. The important thing is that one does not use any of these elusive terms without first defining them, and that, having defined them, one does not use them in any other sense.

(2) Speculation purporting to prove that there must be or that there cannot be a comprehensive process theory has continued for centuries.

Such “metaphysical” preoccupation with “determinism” is, from the viewpoint here taken, irrelevant. Whether and to what extent process theories are successful is entirely a matter of fact. Yet a clear notion of what could be meant by a *comprehensive process theory* is helpful. The way to attain this clarification is to describe the state of affairs that would prevail if we were in possession of such a theory. To do this, consider a finite physical space F , say the office in which I sit while writing this. Now we make two assumptions.

First, we assume that there is a finite class E' of kinds of empirical construct (for example, when the comprehensive theory is “mechanistic,” mass, position, momentum) such that the numerical values at any moment t of *all possible* empirical constructs realized in F can be computed from measurements, at t in F , of constructs of the class E' . The formulas allowing for these computations are sometimes referred to as operational definitions. However, the term is better avoided. For, whenever a construct is defined independently of E' , then the formula for it is not a definition but an empirical law of the cross-sectional type, such as $pv = RT$. *Second*, we assume that there is a process theory applying to *all possible* configurations in F , so that E' is identical with the class E of the theory. A theory satisfying these two conditions would be a comprehensive theory for F .

The space F is ordinarily called a closed system or, in case the “action flow” through its boundary does not vanish, a controlled system. Whenever one fails to find a theory, comprehensive or otherwise, for a piece of space F , one can always renew one’s efforts for another, larger segment of space. As a bare possibility of thought—since it is hard to see how we would ever know it—there could be a deterministic world without a single closed or controlled part in it. More to the point, the notion of closure, like the notion of identity, cannot even be formulated without reference to the success of a theory. After this inductive feature has once been pointed out, one need not burden one’s formulations by always mentioning it. Similarly, the two occurrences of the phrase ‘all possible’ in the foregoing definition refer, not to a vagueness, but to the possibility of refutation of a successful theory by future experience. Such

inductive uncertainty is, according to the view here taken, characteristic of all science and, therefore, always tacitly understood.

If and as long as physics works toward the ideal or under the assumption of an eventual comprehensive theory, we shall say that it is deterministic in the strict sense, or *strictly deterministic*. Instead of speaking of an ideal or an assumption of this kind one could also speak of a frame of reference. These are but different names for the same thing. But then, I believe it is doubtful whether our frame of reference ever was strictly deterministic. Not even Laplace expected that we would eventually be able to predict the outcome of an individual throw of dice under so-called chance conditions. That means that the value of at least one "possible empirical construct" could not be actually computed. There is, however, another, weaker sense, in which physics has been deterministic for a long time. Physics is *deterministic* if, without working toward a comprehensive process theory, it expects to relate all its empirical laws to one and the same process *schema*. Classical physics was deterministic, though not strictly deterministic. Up to the time when the difficulties connected with the ergodic hypothesis suggested certain radical revisions, the schema to which its coordinating definitions related everything was a mechanical model. This is probably the reason why 'mechanistic' and 'deterministic' are sometimes used synonymously. Yet a comprehensive mechanical schema—a comprehensive model—is only one way of realizing determinism in the weaker sense. On the other hand, a comprehensive schema may be "mechanical," at least in the weak sense that it preserves the particle notion, and yet "statistical" in a sense that is not compatible with the process feature of "determinism." All this will be taken up presently. First a few brief comments on the two terms in the phrase 'mechanical model' are in order. In these comments, as throughout this paper, I shall neglect relativistic formulations.

(3) Whether or not it is used as a model, a schema will be called *mechanical* if it is essentially similar to the schema of classical mechanics. Classical mechanics as a theory is, of course, not a model but a so-called phenomenological theory. I shall not bother to repeat what has just been said concerning the fringe of arbitrariness that

surrounds such an expression as 'essentially similar.' The essential feature, as I see it, is the notion of *orbit*. The introduction of, say, an inverse-cube instead of Newton's inverse-square formula or, for that matter, any other change in the form of the Hamiltonian function would thus not be considered an essential dissimilarity. A mechanical schema, then, is a schema whose processes can be interpreted as the successive positions and momentums of points in three-dimensional space. "Momentums," in this formal context, are to be defined as functions of the first derivatives of the "positions" with respect to the parameter "time." A mechanical schema has thus two decisive aspects, the *particle* feature and the *orbit* feature. The particle feature is the occurrence of position-momentum sets; it is not independent of the orbit feature. For wherever there are orbits there are, in this formal sense, also particles. The converse does not hold. Earlier analysts also attributed considerable importance to another feature, the notion of mass and its constancy. More recently, we have come to realize that the only "essential" function of mass was, first, to represent a certain measure of temporal persistence and, second, to serve as identifying tags for the "particles." Further logical analysis has convinced us that such individualization—or spatio-temporal localization, if you please—is already implicit in the notion of orbits.

But there is also the point that in the modern quantum theory the individualization of the particles has been limited in a manner that is natural enough in any theory of particles without orbits. The limitation is represented by the Pauli exclusion principle. For the basic rationale of this principle is that, if we have particles without orbits, the question as to which particle becomes which in two successive temporal cross sections does not even arise. It becomes, as some would rather carelessly say, meaningless. But then, I feel that this particular aspect of the logic of quanta has been sufficiently clarified. So I feel justified in restricting myself to Reichenbach's program as indicated in the opening statement of this paper.

(4) For a proper and tenable analysis of the differences between a *model theory* and a *phenomenological theory* I must refer to my earlier paper.² Here I can only say that a model theory is a theory with a partially coordinated mechanical schema—the model—and that the old kinetic theory before Maxwell and Boltzmann is

still the best paradigm for it. Very roughly speaking, one could say that in explaining the empirical reference of the terms of a model theory one will somewhere find a sentence that begins with 'assume' and contains, later on, the word 'really' in connection with the basic entities of a mechanical schema. For instance, in Newton's corpuscular theory: "Assume that, whenever there is a ray of light there is, really, a beam of moving particles so that. . . ." No such "assumption" is ever encountered in the definitional hierarchy of the empirical constructs.⁵

Perhaps the greatest advantage of what I thus take to be the correct logical description of models is that it keeps one from indulging in a certain kind of dilettante speculation about the "real nature" of the physical universe. Upon such a view, it is said, for instance, that "reality" cannot be both particles and waves. In terms of the structural features of a schema or, rather, as I hope to show, a pattern of schemas, it can. Sometimes the gist of this approach is expressed by saying, in the style of Mach, that the particles of a model are merely computational entities. Unfortunately, this formulation is likely to produce very undesirable psychological reactions against the "subjectivization" of science. Many eminent physicists, Born and Einstein among them, feel very strongly on this point. However that may be, I certainly do not mean to imply that a computational entity in an applicable schema is a fiction or a convention in any foolish or foolishly subjective meaning of these terms.

One of the terminological consequences of the clarification involved is that it restricts the term 'operational' to empirical constructs. If it is to have any specific meaning at all, it makes no sense to use 'operational' or 'operationally meaningless' with reference to a model. All one can require of a model is that it is *somehow* coordinated to the empirical constructs. A model thus coordinated may be redundant in the sense that some of its features are not utilized in the coordinations. To call a redundant feature operationally meaningless makes 'operational' a fashion term, signifying nothing in particular. The worst disadvantage of such loose usage is, perhaps, that it seems to invoke the authority of philosophy to support very specialized physical theories, or conversely. No such support is needed, and none can be given.

Assume, for instance, the classical kinetic theory and the experimental facts in, say, 1890 to have been such that if the former is applied to the latter it becomes a consequence of the former

that no experiment that would allow for the assignment of numerical values to the momentums and positions of individual particles¹⁰ could ever be designed. Such a situation would still not have made it "operationally meaningless" or in any other "philosophical" sense objectionable for the physicists of 1890 to speak of the positions and momentums, exactly or approximately, of individual particles in their model. The question, had it been raised, would have been one of economy, not one of meaning. The bearing of this remark on certain formulations that are now being proposed is obvious. I have already hinted at that in the introduction.

III. Causality, Probability, Statistics

The terms 'causal' and 'causality' serve no particular purpose in the philosophy of science or, for that matter, in analytic philosophy as contrasted with philosophical speculation. If used, they have, so far as I can see, one of the following meanings. (i) Sometimes they refer to the fact that there is, so far as we inductively know, a large number of empirical laws and that we have been rather successful in organizing them by means of theory. Some writers wish to apply the terms only to all or some of the laws and theories of the process type. (ii) Sometimes belief in causality amounts to the assertion of either a deterministic or a strictly deterministic frame of reference. (iii) Sometimes, particularly in recent discussions, causal laws are contrasted with statistical or probability laws, just as causality is contrasted with, or related to, probability. In this section I shall examine some of the questions that are usually considered under the last heading.

(1) At the end of SEC. I, on measurement, I introduced a fiction in order to make a point. The fiction was that our measurements are ideally accurate. The point was that there is no connection between the actual inaccuracy of measurement and the issue of causality *versus* probability as raised in philosophical examinations of the quantum theory. With this in mind, I introduce now, *among empirical constructs*, the distinction

¹⁰ Except, of course, within the range of the empirical constructs; for example, the Cartesian position coordinates of the particles in a 1-cm cube of some gas lie, in the proper coordinate system, between 0 and 1 cm.

between *individual* constructs on the one hand and *statistical* constructs on the other. The distinction itself is as simple as it is familiar; my main concern is with a description that is both sufficiently precise and sufficiently abstract for my purpose.

The present height l of Mr. Smith is an individual construct; the present average height $[l = (1/N) \sum_{i=1}^N l_i]$ of

all American males above 20 years of age is a statistical construct; so are the standard error of this measurement

$[\Delta P = (1/N) \sum_{i=1}^N (l_i - l)^2]$, its higher momentums, and the

so-called distribution of it, either discrete ($f_i, \sum f_i = 1$) or, with the customary assumptions and fictions that allow for

the use of real numbers, continuous $[\varphi(x), \int_{-\infty}^{\infty} \varphi(x) dx = 1]$.

Let me now try to state the thing abstractly.

To determine whether a certain empirical construct is statistical or individual one must trace its definition to what has been called the thing level, or the level of the immediately observable.² The construct is statistical if and only if at some place in this definitional chain a class of measurements has been combined into a new term by a "statistical" procedure such as average formation, computation of a distribution or a standard error. 'Statistical,' as used in the phrase 'statistical procedure,' is an arithmetic term with a precise meaning. The characterization is, therefore, not circular.

Everything that has been said so far refers only to empirical constructs such as length, temperature, pressure. Pressure and temperature in particular are, according to the definition, individual constructs, their interpretation in the kinetic theory notwithstanding. Also, it should be noticed that whether a construct is statistical or individual is a matter of its definition, not of the laws and theories in which it occurs.

(2) An empirical law is the inductive statement of a uniformity; it establishes a relation among several empirical constructs in specified configurations; in the limiting case, the temporal behavior of a single empirical construct. In this respect all laws are alike; they are all—in the first meaning of the term given above—causal laws. They are also all alike in that a single counterinstance that cannot be accounted for otherwise

refutes the law, irrespective of whether it is, as some put it, statistical or causal. The difference lies only in what is to be considered as a counterinstance. In the case of so-called statistical laws, though *practically* the law of large numbers works as well as any of our hypotheses, the *logic* of this decision is rather complex. However, the difficulties that beset it are, in principle, not different from those characteristic of all inductions. Their analysis belongs to the logic of probability and cannot be undertaken in this paper.^{3,11} So I turn now to the description of two types of law. They are both, rather inaccurately, referred to as statistical.

An *ensemble* is a configuration consisting of parts—in the literal, spatial sense of 'part'—that are similar in that they all exemplify one and the same measurable feature. To give a trivial illustration, the objects on my desk form an ensemble with respect to weight and volume. Such features can be used to define statistical constructs which are then characteristic of the ensemble. An empirical law that contains such statistical constructs I shall call an *ensemble law*. As is readily seen, there is nothing particularly statistical about the law itself. It may, for instance, be a process law and, in this sense, "causal." (The schema of radioactive decomposition is of this kind.) On the other hand, an ensemble law may also be a statistical law, in the only sense in which I shall speak of statistical laws. This sense is the same as the one sometimes attached to the expression 'probability law.' However, I shall avoid any formal use of 'probability,' just as I have avoided such use of 'causality.' A *statistical law*, then, is a law, not about a process in a configuration, but about the limiting values of frequencies in an infinite series of similar events in the same configuration. The classical illustration of what I have so abstractly expressed is the repeated throw of a die. A statistical law makes no prediction whatsoever about what is in this case the configuration—in ordinary parlance, the individual event. It is a statement about the "ensemble" formed by the series; but this "ensemble" is merely a verbal, not a real ensemble, if I may so express myself.

¹¹ See also a symposium on probability by Williams, Nagel, Reichenbach, Carnap, Margenau, Bergmann, von Mises and Kaufmann in *Phil. and Phen. Res.* 5-6 (1945).

And, most important, there is no longer a notion of process involved, while we have seen that there may be ensemble process laws. There are, nevertheless, many well-known similarities in the mathematical treatment of statistical laws, ensemble process laws and ensemble statistical laws. The factual assumption that guarantees these relations is the so-called independence of the parts of an ensemble. This, indeed, is what is meant by 'independence' when the term is used in this general area.

Just to practice some of our definitions, I shall mention that what some people seem to mean by 'causal law' is what I would have to call a process law about individual constructs. But this is merely by the way; my main motive in making the distinctions of this section is to call attention to the structure of "nonstatistical" ensemble processes and, also, to some related schemas. So I shall now transfer some of these notions to arithmetic schemas. Only two cases will be considered—the distribution process schema and the statistical schema.

(3) Formally, the fundamental notion is that of a distribution function or, for short, distribution. A distribution is, in the discrete case, a set of non-negative numbers f_i such that $\sum f_i = 1$; in the continuous case, a non-negative function $\varphi(x)$ such that $\int_{-\infty}^{+\infty} \varphi(x) dx = 1$. With the proper mathematical tools the discrete case becomes a specialization of the continuous one. Remembering what has been said in SEC. II, one sees now the possibility of systems whose states are defined as a distribution (in one or several variables) or as an ordered set of such distributions. Such a schema is a special case of a field schema, or, if the process formula is of a certain kind, of a wave schema. If the states of a process schema are distributions, the systems of which these distributions are the states may themselves be given by a set of functions, perhaps even distribution functions. This will turn out to be the case of the quantum-mechanical schema, but of this I shall speak later; for the moment I merely define a *distribution process schema* as a process schema whose states are either single distribution functions or ordered sets of such. Given the system and one of its states—an ordered set of N distributions for the value t_0 of a parameter—the process formula allows for the computation of one and only one ordered set of N distributions for

every value of the parameter. Again, there is nothing particularly "statistical" about this schema; the only feature that could suggest statistics is the occurrence of distributions.

One may be tempted to say that a distribution process schema is always the schema of an ensemble process. However, this is not quite exact. The basic entities that determine the condition of an ensemble are the individual measurements. The resulting distributions are already among the derived or, as one also says, explicitly defined magnitudes of the corresponding schema. In a distribution process schema, on the other hand, the distribution functions are themselves among the basic entities of the calculus. They are, so to speak, not distributions of anything. So all one could say is that the schema of an ensemble process contains a part that, taken by itself, is a distribution process schema. Conversely, the question may be raised whether a distribution process schema can be so supplemented that it becomes the schema of an ensemble process of a certain kind or in a certain manner. Such an attempt may or may not succeed. It will be seen, for instance, that the quantum-mechanical schema cannot be supplemented—in a manner that would be satisfactory in the total context of the theory—to the schema of a "particle ensemble." I shall now explain what I mean by the latter term.

Consider an ensemble whose distributions are those of the positions and momentums of a class of mass points, and make the customary assumption that these distributions are independent within the ensemble. To elucidate the last assumption, assume that $a_i \leq q_i \leq b_i$ for 30 percent of the particles and that $c_j \leq p_j \leq d_j$ for 20 percent of the particles. Then 6 percent of them have both coordinates within the limits indicated. Ensembles whose parts are not individual mass points but configurations of such are essentially of the same kind.

Such a schema determines for every position-momentum set the percentage of part configurations that are "in" this mechanical state.¹² In this

¹² Technically, I should here speak of an arbitrarily small cell of the phase space and of so-called probability density, not of one position-momentum state and its probability. This obvious consequence of working with continuous distributions has nothing whatever to do with inaccuracy or indeterminacy.

respect it preserves the notion of a particle. The process formula, on the other hand, does not connect mechanical states, but distributions of such. Is this the pattern of a mechanical schema, as I have defined the term? Yes and no. It has particles and it has a process, but it has no orbits; its process is, as it were, a statistics of orbits. The schema is undoubtedly an intermediate type of its own; so I suggest for it the intermediate name of a *particle ensemble*. It is interesting for two reasons: first, the schema of the kinetic theory, in its radical interpretation (for example, that of von Mises), is a statistics of particle ensembles; second, the quantum-mechanical schema is nonmechanical in the sense that it cannot even be interpreted as a particle ensemble process within the context of the theory.

(4) Turning to the schema of a statistical law, I consider only the case where states are given by a finite ordered set of N numbers $[x_1^0, \dots, x_N^0]$. The formula is such that it allows for each state the computation of one and only one ordered set of distributions for each value of the parameter t : $\varphi_1(x, t), \dots, \varphi_N(x, t)$. What is here problematical is the use of the term 'process.' The only process feature lies in the typical role of the time parameter. Otherwise the rule states, in ordinary parlance, that given any set of values $[x_1^0, \dots, x_N^0]$ for the zero value of the parameter, the system will in $\varphi_1(a_1, t) \cdot \varphi_2(a_2, t) \dots \varphi_N(a_N, t)$ percent of the cases¹² be in the state $[a_1, a_2, \dots, a_n]$ for the value t of the parameter.

In speaking quite naively of percentages I indicate again that I do not wish to deal with the niceties of the logic of probability. But I shall ask again whether this pattern—call it a *statistical schema*—could be called mechanical, even if its states can be read as position-momentum coordinates. As in the case of the particle ensemble, the answer is: Yes and no. What is preserved, in about the same manner, is the particle notion. What has been discarded is, not only the notion of the orbit but, even more radically, the notion of process. One could say that in the ensemble process the individual process has merely been neglected; at least, it has not been excluded. The statistical law goes further in this respect. This is perhaps the main reason why statistical lawfulness and determinism are often contrasted with each other. For any notion of determinism certainly implies that of process. But then, it has also been seen that this is only one ingredient of an intelligible notion of determinism. The mechanical or particle feature, finally, is not essential to either notion.

(5) If the schema of a theory is of the ensemble or the statistical type (or possibly both), it does not follow that all, or even any, of the empirical laws in the theory are either ensemble processes or statistical. That still depends on how the empirical constructs are coordinated to the entities of the schema. But all this is well known from the early states of the kinetic theory.

(To be concluded.)

The new historians of science, those having obtained better training than we did, will be the best coordinators of scientific education in all its forms, and what is even more important they will constitute the necessary links between our technical barbarians and the well-meaning humanists. They it is who will help us most to integrate our spiritual life—on the one hand by explaining scientific points of view and methods to the humanists, the politicians and the administrators, and on the other by humanizing the men of science and the engineers and reminding them always of the traditions without which our lives, however "efficient," remain ugly and meaningless.—GEORGE SARTON, Isis, Nos. 107–108 (1947), p. 6.

A Simple X-Ray Diffraction Camera

WILLIS C. CAMPBELL

Engineering Experiment Station, University of New Hampshire, Durham, New Hampshire

THE x-ray diffraction camera described here is easily constructed, is readily adjusted to the x-ray beam, requires only short exposures, and affords students in the laboratory an easy introduction to the more common x-ray diffraction methods. By means of simple attachments, it may be used to obtain Laue patterns, back reflections, powder patterns and grazing exposures. An outmoded dental or medical x-ray unit provides enough x-rays for some of the work, and the convenient dental films record diffraction patterns effectively.

As shown in Figs. 1 and 2, a brass plate, $3 \times 3\frac{1}{2} \times \frac{1}{4}$ in., is screwed to the end of a block of wood $2 \times 3 \times 4$ in. A groove cut lengthwise along the top of the wooden block is used to carry a wooden slider which supports a dental x-ray film placed on edge. The slot in which the film rests is made with a hack saw. A hole 1 mm in diameter is drilled through the brass plate $\frac{5}{8}$ in. from its top and midway between its sides, to serve as a pinhole to transmit the x-rays. The camera is clamped to the ring of the ringstand by means of a bolt which passes down through the center of the wooden block and on through a strip of metal extending out under the lower surface of the ring. A wing nut on the lower end of the bolt draws the metal strip against the ring to hold the camera in place. The beam from an x-ray diffraction tube usually makes an angle of about 6° with the horizontal, so it is necessary to provide a support for the camera which may be adjusted to this beam.

Laue (transmission) photographs are taken with the specimen placed over the pinhole on the

side of the brass plate towards the film, as seen in Fig. 2. Plasticine or Scotch Tape is used to support the specimen in this position.

The performance of the camera is shown by the kind of photographs obtained with it. Operating data are given in Tables I and II. The first three patterns of Fig. 3 show symmetry of structure in the samples photographed. The radiating spots in the fourth pattern are characteristic of distorted crystals. Figure 3(b), for sodium chloride, shows more widely spaced diffraction spots than Fig. 3(a), for potassium chloride, because the NaCl structure has smaller lattice¹ constants than the KCl structure.

The first four exposures listed in Table I were made with the radiation from a 150-kv industrial x-ray unit. An outmoded medical unit was used to obtain the patterns in Fig. 4. The aluminum grains in the sample photographed here are arranged in an orderly manner known as preferred orientation, and this condition is evident from the discontinuous rings in the pattern. The mica pattern of Fig. 4(b) shows some symmetry of structure.

Copper radiation was used to obtain the pattern of mica in Fig. 5(a), and the wider separation of the diffraction spots in this case compared to those in Fig. 4(b) is due to the longer wavelength

TABLE I. Data with tungsten radiation.

Figure	Tube voltage (kv)	Tube current (ma)	Specimen to target distance (cm)	Specimen to film distance (cm)	Exposure time (min)
3(a)	60	20	9	1.0	25
3(b)	60	20	9	1.0	25
3(c)	60	20	9	1.5	25
3(d)	60	20	9	2.0	30
4(a)	60	10	6	1.0	10
4(b)	60	10	6	1.0	10

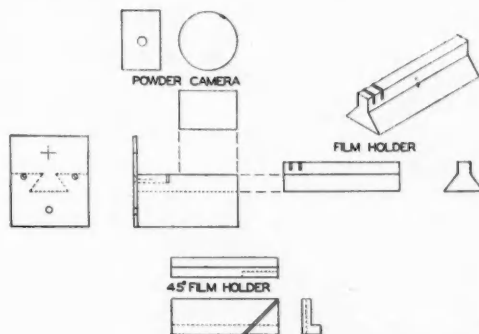


FIG. 1. Drawing of the parts of the camera.

¹Hirst, *X-rays in research and industry* (Chemical Publishing Co., 1943), pp. 15 and 38.

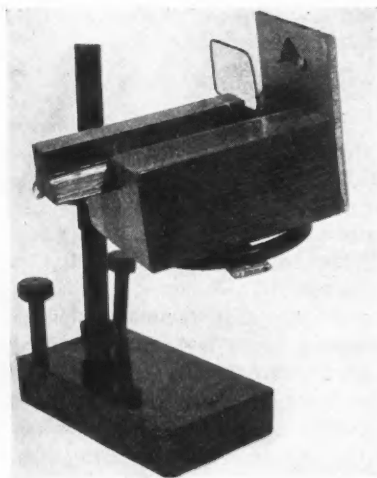


FIG. 2. Camera set up to take Laue photographs.

of the x-rays from the copper target. Figure 5(b) is the pattern of mica while in a bent condition to produce strain. The longer the radiating lines in this type of pattern the greater the strain indicated in the crystals. A lead disk was glued to the center of the film for this exposure so that the direct beam would not obscure the pattern.

The preferred orientation indicated in the pattern of cold-rolled aluminum [Fig. 6(a)] disappears when recrystallization takes place during the heating and cooling processes [Fig. 6(b)].

Cold-rolled copper shows preferred orientation in its diffraction pattern [Fig. 7(a)], while electrolytic copper [Fig. 7(b)] shows random orientation (continuous rings).

For back-reflection photographs the specimen is mounted with Plasticine on the end of the shaft of an electric motor (Fig. 8). The film is placed in a slot which makes an angle of 45° with the x-ray beam, and in this position all back reflections will be recorded. The film is too small to catch all of the reflection when it is mounted on the brass plate.

Rotating a specimen during exposure gives smooth lines [Fig. 9(b)] so that their distances from the main beam may be measured accurately. This is a nondestructive test often used on a sample that is too thick to allow the x-ray beam to pass through it.

Aluminum, silver and copper have face-centered cubic lattices, but the spacings and densities of their diffraction lines are seen to be different (Figs. 9 and 10). Zinc has a hexagonal close-packed lattice, and tungsten a body-centered cubic lattice. Each structure has its own distinctive pattern. If the developed film is replaced in the camera in its original position, a pair of dividers may be used to measure the distance from the pinhole to any diffraction line. With this value and the distance of the specimen from the pinhole laid off to scale on polar

TABLE II. Data with copper radiation; standard x-ray diffraction tube.

Figure	Tube voltage (kv)	Tube current (ma)	Specimen to target distance (cm)	Specimen to film distance (cm)	Exposure time (min)
5(a)	40	10	3.0	1.00	2
5(b)	40	10	3.0	1.00	2
6(a)	25	5	3.0	1.00	5
6(b)	25	5	3.0	1.00	5
7(a)	35	10	3.0	1.00	5
7(b)	35	10	3.0	1.00	5
9(a)	40	10	4.5	1.75	10
9(b)	40	10	4.5	1.75	10
10	40	10	4.5	1.75	10
12	40	10	4.5	2.00	10
14	40	10	7.5	2.75	5

coordinate paper, the Bragg angle may be read directly. The line may then be indexed using the wavelength of the x-ray beam and the lattice

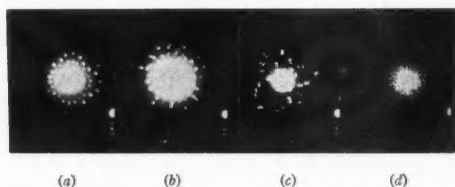


FIG. 3. Laue photographs: (a) KCl crystals; (b) NaCl crystals; (c) calcite crystals; (d) iron sample.

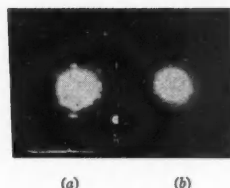


FIG. 4. Laue photographs: (a) cold-rolled aluminum; (b) mica.

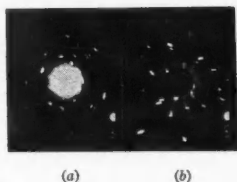


FIG. 5. Laue photographs: (a) mica; (b) bent mica.

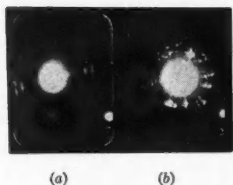


FIG. 6. Laue photographs: (a) cold-rolled aluminum; (b) aluminum after heating.

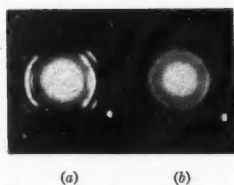


FIG. 7. Laue photographs: (a) cold-rolled copper; (b) electrolytic copper.

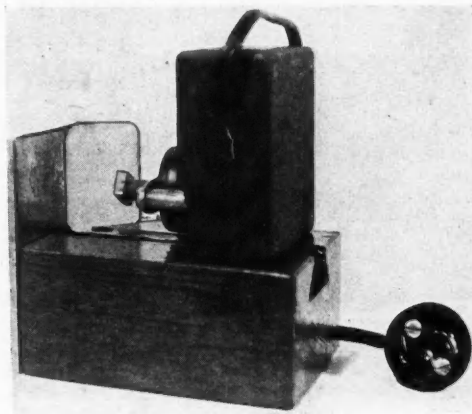


FIG. 8. Camera set up to take back-reflection photographs.

With the x-ray beam perpendicular to the rolling direction, the first pattern in Fig. 12 shows arcs; but with the x-ray beam parallel to the rolling direction, the second pattern shows continuous rings. Since the third and fourth patterns are similar there is no evidence of cold work in the sample examined.

A 35-mm film container (metal) is used in the camera set up to take powder photographs (Fig. 13). The inside of this metal can is lined with sheets of lead discarded from dental film packs. Two holes $\frac{1}{4}$ in. in diameter are drilled

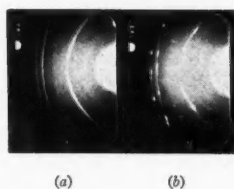


FIG. 9. Back-reflection exposures: (a) aluminum, not rotated; (b) aluminum, rotated.

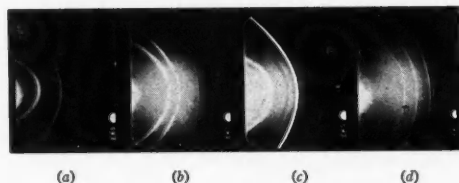


FIG. 10. Back-reflection exposures: (a) silver; (b) copper; (c) tungsten; (d) zinc.

constants of the substance, in the proper formulas.²

A right-angled block of wood is placed on the camera (Fig. 11) to support the film in the proper position to record grazing exposures. The slot in which the film rests makes an angle of 45° with the x-ray beam. A small mound of Plasticene on the wooden slider supports the specimen, which is adjusted to allow the x-ray beam to strike its upper surface at an angle of about 10° . Part of the direct x-ray beam passes over the top edge of the specimen without being deflected, and the rest either passes through the thinner part of the specimen or is reflected from the atoms near its surface. Some of the effects of cold work on metals may be shown by this method of x-ray diffraction.³

² Barrett, *Structure of metals* (McGraw-Hill, 1942), p. 76.

³ Davey, *A study of crystal structure and its applications* (McGraw-Hill, 1934), p. 514.

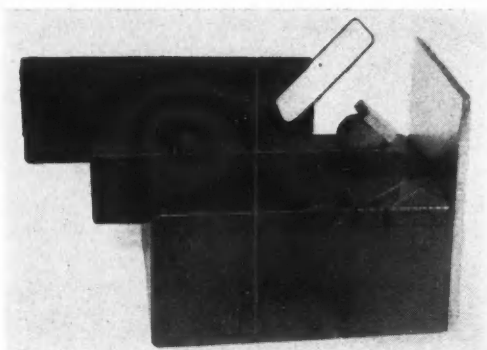


FIG. 11. Camera set up to take grazing photographs.

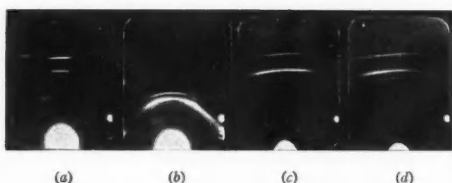


FIG. 12. Grazing exposures: (a) aluminum, x-ray beam perpendicular to rolling direction; (b) x-ray beam parallel to rolling direction; (c) copper sample; (d) same copper sample turned through 90° .

through the opposite sides of the can and midway between its top and bottom, to serve as inlet and outlet, respectively, for the x-ray beam. The can is set on Plasticine on top of the wooden block. The brass plate on the camera is reversed from its former position, and a hole is drilled through it in line with the two holes in the can. This hole in the plate takes a collimating tube which has $\frac{1}{4}$ -mm lead pinholes. The parallel rays obtained with the collimating tube give sharper diffraction lines, with very little fogging of the film. A disk of heavy cardboard, with a diameter slightly less than that of the metal container, is glued to the bottom of the container. The film fits tightly between the rim of the cardboard and the inner surface of the can. A small mound of Plasticine on the cardboard disk supports the specimen in a vertical position in the center of the container. Black paper is glued over the holes in the camera. The specimen for the powder pattern of copper (Fig. 14) is prepared from powdered copper which had passed through a 200-mesh screen. A little of this powder is mixed with household glue to make a dough-like paste, which is then rolled between

two ground-glass plates to form a cylindrical specimen about $\frac{3}{4}$ mm in diameter. After the glue has hardened the specimen is ready to be mounted in the camera. A strip of no-screen x-ray film, 1×6 in., with a hole punched in its center, is used to record the diffraction pattern. Before loading the camera it is lined up so that the specimen is in the center of the x-ray beam as seen on a fluorescent screen. (Lead glass is used to protect the operator from the x-rays while making this adjustment.) After putting the film in the can of the camera, a final check is made on the x-ray beam as the exposure is started.

The powder method of x-ray diffraction is the most useful one. The random orientation of the many finely divided grains in the powder permits more reflections than are possible by any other method. These reflections are needed to identify the kind and amount of material in the analyzed substance.

Figure 14 is typical of the patterns obtained with a powder camera. The diffraction lines nearer the hole in the center of the film are due to the rays transmitted through the specimen, while the other lines are due to the rays reflected from the specimen (back reflections). The identification of a substance is completed if its pattern matches that of a known substance. If patterns

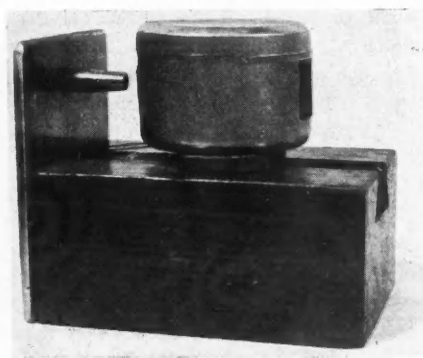


FIG. 13. Camera set up to take powder photographs.

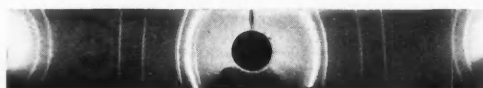


FIG. 14. Powder pattern of copper.

of known substances are not available, mathematical computations are made and the results compared with (Hanawalt)⁴ index cards which cover a wide range of materials.

⁴ Sproull, *X-rays in practice* (McGraw-Hill, 1946), p. 423.

X-ray diffraction is solving more and more of the problems of modern science. A simple x-ray diffraction camera may be used in place of more expensive equipment as an introduction to this interesting and useful subject.

The Problem of Impact Analyzed by Tensor Methods: Classical Dynamics

MARCELO ALONSO*

Universidad de la Habana, Habana, Cuba

IT is well known that vector methods are now widely used in physics because of the great simplification they introduce in the analytic treatment of many problems. Our purpose is to apply vector analysis to the problem of impact and thus to show once more the power of this method of attack on physical problems. We shall also make free use of dyadics. For the sake of clearness we shall explain first some fundamental facts about dyadics. In this paper we use only classical, or nonrelativistic, dynamics; but we hope to complete the treatment in a second paper using relativistic dynamics. We also assume that the colliding bodies are perfectly smooth and nonrotating spheres. This hypothesis is introduced in order to simplify the calculations and because in that case we do not need to consider a possible exchange of energy of rotation between the two colliding bodies.

1. Fundamental Facts about Dyadics¹

Consider two vectors \mathbf{a} and \mathbf{b} . The expression \mathbf{ab} is called a *dyad* and must be considered as an operator. Operating scalarly with \mathbf{ab} as a prefactor on another vector \mathbf{V} , that is, multiplying scalarly \mathbf{ab} and \mathbf{V} , we get²

$$\mathbf{ab} \cdot \mathbf{V} = \mathbf{a}(\mathbf{b} \cdot \mathbf{V}).$$

This means that the operation of \mathbf{ab} on \mathbf{V} transforms \mathbf{V} into another vector, in the direction of \mathbf{a} , whose magnitude is that of \mathbf{a} multiplied by $\mathbf{b} \cdot \mathbf{V}$. Evidently $\mathbf{V} \cdot \mathbf{ab} = (\mathbf{V} \cdot \mathbf{a})\mathbf{b}$ is different from $\mathbf{ab} \cdot \mathbf{V}$ unless \mathbf{a} and \mathbf{b} are parallel.

* Now at Instituto del Vedado, Habana, Cuba.

¹ For the general theory of dyadics see W. Gibbs and E. B. Wilson, *Vector analysis* (Yale Univ. Press), chap. V; or L. Page, *Introduction to theoretical physics* (Van Nostrand, ed. 2), art. 18.

² There are of course other possible multiplications or operations between \mathbf{ab} and \mathbf{V} .

A *dyadic* is a sum of dyads: $\mathbf{\Psi} = \mathbf{ab} + \mathbf{cd} + \dots$. Dyadics, vectors and scalars are special cases of quantities of a more general type, called *tensors*. A dyadic is a tensor of the second rank; a vector, a tensor of the first rank; and a scalar, a tensor of zeroth rank.

There is one very important dyadic, such that it leaves unchanged any vector on which it operates. It is called the *idemfactor*, or unit tensor. Its expression in terms of three orthogonal unit vectors \mathbf{i} , \mathbf{j} , \mathbf{k} parallel to a set of rectangular axes is

$$\mathbf{I} = \mathbf{ii} + \mathbf{jj} + \mathbf{kk}.$$

Clearly,

$$\mathbf{I} \cdot \mathbf{V} = \mathbf{ii} \cdot \mathbf{V} + \mathbf{jj} \cdot \mathbf{V} + \mathbf{kk} \cdot \mathbf{V} = iV_x + jV_y + kV_z = \mathbf{V}.$$

Also, $\mathbf{V} \cdot \mathbf{I} = \mathbf{V}$.

Consider again the vector $\mathbf{V} = iV_x + jV_y + kV_z$. Its projection on the xy plane, perpendicular to the unit vector \mathbf{k} , is $\mathbf{V}' = iV_x + jV_y$. Using the preceding relations, we obtain

$$\mathbf{V}' = \mathbf{ii} \cdot \mathbf{V} + \mathbf{jj} \cdot \mathbf{V} = (\mathbf{ii} + \mathbf{jj}) \cdot \mathbf{V} = (\mathbf{I} - \mathbf{kk}) \cdot \mathbf{V}.$$

Therefore the dyadic $\mathbf{ii} + \mathbf{jj} = \mathbf{I} - \mathbf{kk}$ transforms any vector \mathbf{V} into its projection on the xy plane, perpendicular to \mathbf{k} .

In general, if

$$\mathbf{P} = \mathbf{I} - \mathbf{nn},$$

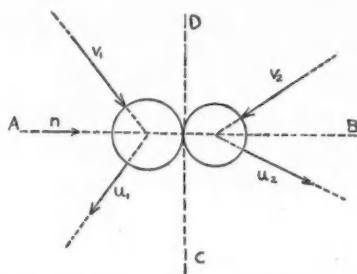
then $\mathbf{P} \cdot \mathbf{V}$ is the projection of \mathbf{V} on a plane perpendicular to the unit vector \mathbf{n} . We call \mathbf{P} a *projection operator*.

2. General Equations for Impact

Consider two bodies of masses m_1 and m_2 . Their velocities before impact will be designated by \mathbf{v}_1 and \mathbf{v}_2 , respectively, and those after impact by \mathbf{u}_1 and \mathbf{u}_2 . Since during the impact we must consider only internal forces of the system, we conclude that the total linear momentum is conserved; that is,

$$m_1 \mathbf{u}_1 + m_2 \mathbf{u}_2 = m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2. \quad (1)$$

The second equation required for a complete analytic description of impact is Newton's empirical law of restitution. We call the line AB

FIG. 1. Oblique impact; AB is the line of impact.

(Fig. 1) perpendicular to the common tangent plane CD at the instant of impact the *line of impact*, and we designate its direction by the unit vector \mathbf{n} . Then Newton's law of restitution states that the component of the relative velocity of the two bodies in the direction of the line of impact, after impact, is proportional to the component of the relative velocity in the same direction, before impact, but has the opposite direction. Analytically,

$$\mathbf{n} \cdot (\mathbf{u}_1 - \mathbf{u}_2) = -e \mathbf{n} \cdot (\mathbf{v}_1 - \mathbf{v}_2), \quad (2)$$

where e is a constant of proportionality called the *coefficient of restitution*, or *collision coefficient*. In general, e may have any value, either positive or negative, depending on the type of interaction between the colliding bodies. If $|e| < 1$ we have a *collision of the first kind*, and if $|e| > 1$ we have a *collision of the second kind*. The reason for this classification will be explained later in connection with the exchange of energy during impact. However, if the impact is of elastic nature—that is, if only elastic forces act during the collision—then e is a measure of the elastic properties of the materials composing the colliding bodies and must lie between 0 and 1. In that case we have *elastic impact*. If $e = 1$ then the impact is *perfectly elastic*; but if $e = 0$ the impact is *inelastic*, or *plastic*, though the latter term is not widely used at present. The reason for this second name is that we have inelastic (or plastic) impact when the stresses developed during the impact are larger than those corresponding to the plastic limit³ of the materials and the bodies remain deformed after impact.

³ R. S. Lindsay, *Physical mechanics* (Van Nostrand), p. 320.

During the impact the components of the velocities in any direction perpendicular to the line of impact, that is, parallel to plane CD , do not change. (This is true only if there is no rotation of the colliding spheres.) So if \mathbf{P} is the projection operator on plane CD , we must have

$$\mathbf{P} \cdot \mathbf{u}_1 = \mathbf{P} \cdot \mathbf{v}_1, \quad \mathbf{P} \cdot \mathbf{u}_2 = \mathbf{P} \cdot \mathbf{v}_2.$$

Subtracting, we get

$$\mathbf{P} \cdot (\mathbf{u}_1 - \mathbf{u}_2) = \mathbf{P} \cdot (\mathbf{v}_1 - \mathbf{v}_2). \quad (3)$$

Multiplying Eq. (2) by \mathbf{n} , adding the result to Eq. (3) and taking into account the fact that $\mathbf{nn} + \mathbf{P}$ is the unit tensor or idemfactor \mathbf{I} , we get

$$\mathbf{u}_1 - \mathbf{u}_2 = (\mathbf{P} - e \mathbf{nn}) \cdot (\mathbf{v}_1 - \mathbf{v}_2),$$

or, adding and subtracting \mathbf{nn} inside the parenthesis,

$$\mathbf{u}_1 - \mathbf{u}_2 = \mathbf{v}_1 - \mathbf{v}_2 - (1 + e) \mathbf{nn} \cdot (\mathbf{v}_1 - \mathbf{v}_2). \quad (4)$$

This equation may be considered as Newton's law in tensor form. Though apparently more general than Eq. (2), it is in fact more restricted because it includes Eq. (3), which, as has been stated before, is valid only in the case of smooth nonrotating spheres.

Equations (1) and (4) are sufficient to solve the problem. To obtain \mathbf{u}_1 we multiply Eq. (4) by m_2 and add to Eq. (1). Then

$$(m_1 + m_2) \mathbf{u}_1 = (m_1 + m_2) \mathbf{v}_1 - m_2 (1 + e) \mathbf{nn} \cdot (\mathbf{v}_1 - \mathbf{v}_2).$$

Therefore,

$$\mathbf{u}_1 = \mathbf{v}_1 - \frac{m_2}{m_1 + m_2} (1 + e) \mathbf{nn} \cdot (\mathbf{v}_1 - \mathbf{v}_2). \quad (5)$$

Similarly, to obtain \mathbf{u}_2 multiply Eq. (4) by m_1 and subtract it from Eq. (1):

$$(m_1 + m_2) \mathbf{u}_2 = (m_1 + m_2) \mathbf{v}_2 + m_1 (1 + e) \mathbf{nn} \cdot (\mathbf{v}_1 - \mathbf{v}_2).$$

Therefore,

$$\mathbf{u}_2 = \mathbf{v}_2 + \frac{m_1}{m_1 + m_2} (1 + e) \mathbf{nn} \cdot (\mathbf{v}_1 - \mathbf{v}_2). \quad (6)$$

Equations (5) and (6) give the velocities after impact in terms of the velocities before impact and are therefore the solution of the problem of impact. We next consider some special cases.

3. Perfectly Elastic Impact

If the impact is perfectly elastic, $e=1$, and Eqs. (5) and (6) become

$$\mathbf{u}_1 = \mathbf{v}_1 - \frac{2m_2}{m_1 + m_2} \mathbf{nn} \cdot (\mathbf{v}_1 - \mathbf{v}_2), \quad (7)$$

$$\mathbf{u}_2 = \mathbf{v}_2 + \frac{2m_1}{m_1 + m_2} \mathbf{nn} \cdot (\mathbf{v}_1 - \mathbf{v}_2); \quad (8)$$

or, if $m_1 = m_2$,

$$\mathbf{u}_1 = \mathbf{v}_1 - \mathbf{nn} \cdot (\mathbf{v}_1 - \mathbf{v}_2), \quad (9)$$

$$\mathbf{u}_2 = \mathbf{v}_2 + \mathbf{nn} \cdot (\mathbf{v}_1 - \mathbf{v}_2). \quad (10)$$

If \mathbf{P} is the projection operator on the plane CD perpendicular to the line of impact, $\mathbf{V} = \mathbf{I} \cdot \mathbf{V} = (\mathbf{P} + \mathbf{nn}) \cdot \mathbf{V} = \mathbf{P} \cdot \mathbf{V} + \mathbf{nn} \cdot \mathbf{V}$, and Eqs. (9) and (10) may be written in the form

$$\mathbf{u}_1 = \mathbf{P} \cdot \mathbf{v}_1 + \mathbf{nn} \cdot \mathbf{v}_2, \quad (11)$$

$$\mathbf{u}_2 = \mathbf{P} \cdot \mathbf{v}_2 + \mathbf{nn} \cdot \mathbf{v}_1. \quad (12)$$

This shows that the velocity components perpendicular to the line of impact remain the same while the components parallel to it are interchanged, as is shown in Fig. 2.

If the impact is direct, that is, if \mathbf{v}_1 and \mathbf{v}_2 are parallel to the line of impact, $\mathbf{P} \cdot \mathbf{v}_1 = \mathbf{P} \cdot \mathbf{v}_2 = 0$, and Eqs. (11) and (12) may be written in scalar form,

$$u_1 = v_2, \quad u_2 = v_1,$$

showing that the two bodies interchange their velocities.

4. Inelastic, or Plastic, Impact

If the impact is inelastic, or plastic, $e=0$, and Eqs. (5) and (6) become

$$\mathbf{u}_1 = \mathbf{v}_1 - \frac{m_2}{m_1 + m_2} \mathbf{nn} \cdot (\mathbf{v}_1 - \mathbf{v}_2), \quad (13)$$

$$\mathbf{u}_2 = \mathbf{v}_2 + \frac{m_1}{m_1 + m_2} \mathbf{nn} \cdot (\mathbf{v}_1 - \mathbf{v}_2). \quad (14)$$

These two equations may also be written in the form

$$\mathbf{u}_1 = \mathbf{P} \cdot \mathbf{v}_1 + \mathbf{nn} \cdot (m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2) / (m_1 + m_2), \quad (15)$$

$$\mathbf{u}_2 = \mathbf{P} \cdot \mathbf{v}_2 + \mathbf{nn} \cdot (m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2) / (m_1 + m_2), \quad (16)$$

showing that after impact the components of the

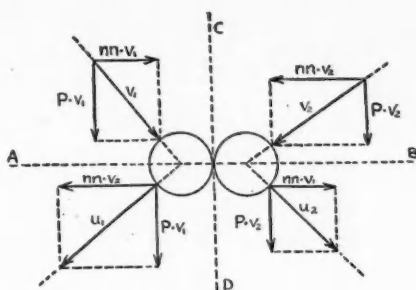


FIG. 2. Interchange of velocities parallel to the line of impact in perfectly elastic impact.

velocities in the direction of the line of impact are the same. Of course, this result could be obtained directly from Eq. (2) by putting $e=0$. The result is $\mathbf{n} \cdot \mathbf{u}_1 = \mathbf{n} \cdot \mathbf{u}_2$, which is equivalent to the preceding statement.

If the impact is direct, $\mathbf{P} \cdot \mathbf{v}_1 = \mathbf{P} \cdot \mathbf{v}_2 = 0$, and Eqs. (15) and (16) may be written in scalar form,

$$u_1 = u_2 = (m_1 u_1 + m_2 u_2) / (m_1 + m_2), \quad (17)$$

showing that after impact the spheres have the same velocity and move as a single body. In general, both have the same velocity after impact if $\mathbf{P} \cdot \mathbf{v}_1 = \mathbf{P} \cdot \mathbf{v}_2$. In that case they also continue like a single body.

5. Dissipation of Kinetic Energy

In general, kinetic energy E is not conserved in the impact; part of it is transformed into some other form of energy. Our objective now is to compute the amount of energy transformed.

Adding the square of Eq. (4) multiplied by $m_1 m_2$ to the square of Eq. (1), we obtain

$$\begin{aligned} (m_1^2 + m_1 m_2) u_1^2 + (m_2^2 + m_1 m_2) u_2^2 \\ = (m_1 \mathbf{v}_1 + m_2 \mathbf{v}_2)^2 \\ + m_1 m_2 [\mathbf{v}_1 - \mathbf{v}_2 - (1+e) \mathbf{nn} \cdot (\mathbf{v}_1 - \mathbf{v}_2)]^2. \end{aligned}$$

Therefore,

$$\begin{aligned} (m_1 + m_2) (m_1 u_1^2 + m_2 u_2^2) \\ = (m_1 + m_2) (m_1 v_1^2 + m_2 v_2^2) \\ - (1-e^2) [\mathbf{n} \cdot (\mathbf{v}_1 - \mathbf{v}_2)]^2 m_1 m_2. \end{aligned}$$

Division by $2(m_1 + m_2)$ gives

$$\begin{aligned} \frac{1}{2} m_1 u_1^2 + \frac{1}{2} m_2 u_2^2 = \frac{1}{2} m_1 v_1^2 + \frac{1}{2} m_2 v_2^2 \\ - \frac{1}{2} (1-e^2) \frac{m_1 m_2}{m_1 + m_2} [\mathbf{n} \cdot (\mathbf{v}_1 - \mathbf{v}_2)]^2, \end{aligned}$$

or

$$E_{\text{after}} = E_{\text{before}} - \frac{1}{2}(1 - e^2) \times \frac{m_1 m_2}{m_1 + m_2} [\mathbf{n} \cdot (\mathbf{v}_1 - \mathbf{v}_2)]^2. \quad (18)$$

So the energy transformed is

$$E_{\text{before}} - E_{\text{after}} = \frac{1}{2}(1 - e^2) \times \frac{m_1 m_2}{m_1 + m_2} [\mathbf{n} \cdot (\mathbf{v}_1 - \mathbf{v}_2)]^2. \quad (19)$$

Kinetic energy is conserved only in the perfectly elastic impact, because then $e = 1$. The maximum dissipation corresponds to the inelastic, or plastic, impact ($e = 0$). It is interesting to observe that the amount of energy dissipated is determined only by the component of the relative velocity parallel to the line of impact, before impact, and is independent of the sign of e .

In the collisions of the first kind we always have $e^2 < 1$; then $E_{\text{after}} < E_{\text{before}}$, and part of the energy of the system is transferred to some external system. This transferred energy usually appears as sound, but in some cases it may be

stored as internal motion of the molecules of the bodies, that is, as heat. In the case of atomic collisions the kinetic energy lost may be used to raise the atoms to excited states. For example, in the case of the impact of a fast electron and an atom or molecule the result may be a slow electron leaving the atom or molecule in an excited quantum state.

On the other hand, in collisions of the second kind $e^2 > 1$ and then $E_{\text{after}} > E_{\text{before}}$, showing that there is a gain of kinetic energy. This may result, for example, when a slow electron collides with an atom or molecule in an excited quantum state. The result in general is a fast electron leaving the atomic system in a lower or in the ground state.

Summarizing this section we may say that in collisions of the first kind the system loses kinetic energy that is transferred to a second system or stored as some form of internal energy; but in collisions of the second kind there is a transfer of some kind of internal energy (or perhaps of kinetic energy of rotation) into kinetic energy of translation.⁴

⁴ See, for example, A. E. Ruark and H. C. Urey, *Atoms, molecules and quanta* (McGraw-Hill), chap. XIV.

I believe that the university laboratories will continue to be the great producers of revolutionary discoveries and of that knowledge of the materials, forces and phenomena of nature which is the basis of industrial research operations. There is a logical reason to back up this belief, as well as the weight of experience. This reason is the greater freedom and range of interest which is characteristic of a university as compared with an industrial organization. For the latter is intellectually circumscribed by the scope of its particular business, and its thought and work are aimed at the solution of its own problems. This is a favorable situation for work to improve a product, or to find a new product, or to cheapen a method of manufacture. It is not so favorable to the development of original and unexpected new ideas, for these almost always appear useless at first—mere intellectual curiosities—and few business organizations can afford the luxury of paying wages for pursuing undirected intellectual curiosities.

Some of our great industrial laboratories of today are very liberal in their policies and produce splendid scientific work as well as engineering development. But the limitations which I have mentioned are always there to a greater or lesser degree—usually greater. . . .—K. T. COMPTON.

Definition of Electric Charge Derived from Simple Quantitative Experiments with Pith Balls

AUSTIN J. O'LEARY

The City College of New York, New York 10, New York

A COMMON method of introducing electric charge goes something like this: after the usual preliminary demonstrations and discussion, we introduce the inverse-square law of electric force, citing the Coulomb experiment; then we simply state that the electric force between two particles is directly proportional to the product of their charges, define unit charge and let it go at that. How does one know that each electrified particle may be characterized by any such single-valued quantity? The confirmation is so indirect, we generally do not call attention to it when we come to measurements that apply; sometimes we do not even raise the question of confirmation in the first place. How can our presentation be improved?

I wish to outline a procedure in which the definition of charge is derived from simple quantitative experiments with pith balls which the students can easily understand and which can be performed, at least qualitatively, in the lecture room; quantitative results may be taken from this paper. The idea is to show how an original investigator might arrive at the quantity *charge*. Students may be given valuable experience in the thought processes that guide experimental procedure by having them share in an analysis of each step; the physical arrangement and the operations performed are simple enough so that attention may be concentrated on plan and interpretation of the measurements.

General Arrangement

A number of uniform balls about 4 mm in diameter were turned off on a lathe from a stick of pith. The balls were smoothly coated with gold foil; all excess foil was rubbed off. Each was fastened to a fine nylon thread. The balls in finished form were carefully weighed on an analytic balance. Four balls very nearly equal in weight (about 8 mg) were selected from the group. The suspension thread of each was passed into a hollow support rod by way of a No. 1-80

drill hole at the tip of the rod and was fastened to a screw that could be raised or lowered in the rod to adjust the length of the suspension. The rods were made with a collar on top and a section of reduced diameter at the bottom so that they could be quickly slipped in or out of snug-fitting tubular holders with a lengthwise slot cut out to allow the smaller diameter of the rod to pass through freely.

At the beginning of each series of observations, the balls were similarly charged by contact with an electrified glass rod or rubber rod. Whichever two balls were to be observed were placed in a cubical enclosure with glass sides to eliminate motion due to air currents; we used a spectrograph case that happened to be at hand. A fine thread was fastened in a vertical position toward the back of the enclosure as a reference line. Each configuration to be measured was photographed on 35-mm film to permit measurements to be made at leisure through projection with an overall magnification of $10\times$. Figure 1 is a photograph of two balls.

The mutual force F between two such balls is very closely equal to $\bar{m}g(d-s)/2h$, where \bar{m} is their mean mass, d the distance between centers, s the distance between tips of the supporting rods and h the vertical height from either ball to either tip. The distances were measured with appropriate refinement.

Inverse-Square Law

We first confine our attention to two pith balls. Starting with the two holders side by side, one holder is kept fixed and the other is moved sideways about 10 cm by means of a leadscrew rotated between pictures, one or two turns at a time, then back to the starting point with the same intermittent motion. Two pictures were taken in short succession at each position to provide a check on the state of rest of the balls.

As the distance between balls is changed, one may easily see that the larger the distance, the more nearly vertical are the suspension threads;

that is, as d increases, F decreases, and *vice versa*. An experienced investigator immediately wonders whether F may not be proportional to some negative power of d . If $F \propto d^n$ (n negative), then $\log F \propto n \log d$. In Fig. 2, $\log F$ is plotted as a function of $\log d$ for corresponding values of F and d from a series of photographs taken in 3 min. The graph is a straight line with slope $n = -2.00$. Thus, our first attempt to find a correlation between F and d is successful; we find that $F \propto 1/d^2$. More specifically, Fd^2 is independent of d for a given pair of balls. Comparing points in Fig. 2 from time $t=0$ to $t=1.5$ min, surrounded by circles, with points from $t=1.5$ min to $t=3$ min, surrounded by squares, we may see that Fd^2 does decrease slightly in the course of time. From photographs taken over a long period of time, it can be shown that $\log Fd^2$ decreases linearly with time. The median line in Fig. 2, therefore, characterizes the electrified state of the pair of pith balls at the instant $t=1.5$ min. In effect, we



FIG. 1. Photograph of two pith balls, selected to indicate the least favorable conditions for measurement ($Fd^2=21$ units, where 1 unit= 10^{-10} newton m^2 ; the values of Fd^2 measured in different instances ranged from this minimum up to 111 units). On the original negative the suspension threads are clearly visible throughout their full length.

have reduced performance of the experiment to a single instant.

Comparison of Four Electrified Pith Balls; Definition of Charge

We now set out to investigate the values of Fd^2 for the six possible combinations in which four pith balls may be brought successively near each other in pairs. Starting with a given pair in place in the glass case, say No. 1 and No. 2, two photographs are taken in short succession; then a movable half of the cover is swung to one side about a vertical axis, ball No. 2 is replaced by No. 3, the cover is swung back, photographs are taken when the balls have come to rest, and the time is recorded; and so on until all six comparisons have been made four times. After some practice, two of us were able to take a complete series of photographs in 8 min. The values of Fd^2 for one series are plotted as a function of t in Fig. 3. A straight line fits each set of four points rather well.¹ Hence, from the graphs in Fig. 3, we can get simultaneous values of Fd^2 for the six different combinations. We look for some correlation among them.

What kind of correlation would it seem most hopeful to look for? There are no mechanical differences among the pith balls. One's first thought then is of some new (electric) quantity to distinguish one pith ball from another—some measure we may think of as representing "amount of electrification." Can we find four quantities q_1, q_2, q_3 and q_4 to characterize the pith balls of corresponding number such that, for any two balls x and y , some combination of q_x and q_y is proportional to $(Fd^2)_{x,y}$? Can we, for example, find four quantities q_i such that $q_x + q_y = \epsilon (Fd^2)_{x,y}$, where ϵ is a proportionality factor? Since experience leads us to expect $(Fd^2)_{x,y}$ to approach zero as either q_x or q_y approaches zero, $q_x + q_y$ would seem to be a most unlikely combination to set equal to $(Fd^2)_{x,y}$; anyway, quite apart from what we might expect, we may soon find by trial that the values of Fd^2 in Fig. 3 do not satisfy this relation. Then how about a product, $q_x q_y$? Can

¹ Actually, Fd^2 for each pair of pith balls is an exponential function of t . However, for the small percentage changes occurring during the short period of observation in this instance, the observed relation between Fd^2 and t is linear within the limits of error of the measurements.

we find four quantities q_i such that

$$q_x q_y = \epsilon (Fd^2)_{x,y} \quad (1)$$

for any pair?

Inserting the values of Fd^2 from Fig. 3, for $t=2$ min, Eq. (1) yields

$$q_2 q_3 \epsilon^{-1} = 110.2, \quad (2) \quad q_2 q_4 \epsilon^{-1} = 39.3, \quad (5)$$

$$q_2 q_4 \epsilon^{-1} = 72.2, \quad (3) \quad q_1 q_3 \epsilon^{-1} = 34.7, \quad (6)$$

$$q_1 q_2 \epsilon^{-1} = 62.3, \quad (4) \quad q_1 q_4 \epsilon^{-1} = 22.7, \quad (7)$$

all expressed in terms of 10^{-10} newton m^2 . Solving Eqs. (2), (3) and (5), for example, we get $q_2 \epsilon^{-1} = 14.23$ units, $q_3 \epsilon^{-1} = 7.76$ units, $q_4 \epsilon^{-1} = 5.08$ units, where the unit is 10^{-6} newton m^2 . Substituting these values in Eqs. (4), (6) and (7), we get for $q_1 \epsilon^{-1}$ the values 4.38, 4.48 and 4.47 units. Agreement among these three values for $q_1 \epsilon^{-1}$ is good enough to give us some confidence in Eq. (1).

However, we get a better picture of the validity of Eq. (1) in another way. We first solve the six trial relations for average values of all four unknowns. If L_i denote $\log q_i \epsilon^{-1}$, Eqs. (2) to (7) become

$$L_2 + L_3 = 2.0422, \quad (2)' \quad L_2 + L_4 = 1.5944, \quad (5)'$$

$$L_2 + L_4 = 1.8585, \quad (3)' \quad L_1 + L_3 = 1.5403, \quad (6)'$$

$$L_1 + L_2 = 1.7945, \quad (4)' \quad L_1 + L_4 = 1.3560. \quad (7)'$$

Adding the six equations and dividing by 3, we get $\Sigma L_i = 3.3953$. Adding Eqs. (4)', (6)' and (7)', we get $2L_1 + \Sigma L_i = 4.6908$, $L_1 = 0.6478$; in like manner, we get $L_2 = 1.1500$, $L_3 = 0.8908$, $L_4 = 0.7068$, which gives us

$$q_1 \epsilon^{-1} = 4.44 \text{ units,}$$

$$q_3 \epsilon^{-1} = 7.78 \text{ units,}$$

$$q_2 \epsilon^{-1} = 14.13 \text{ units,}$$

$$q_4 \epsilon^{-1} = 5.09 \text{ units.}$$

Finally, we test Eq. (1) by comparing experimental and calculated values of Fd^2 . Multiplying the preceding average values of $q_i \epsilon^{-1}$ in pairs, we get

$$109.9, 71.9, 62.8, 39.6, 34.6, 22.6,$$

in good agreement with the corresponding experimental values in Eqs. (2) to (7). At time $t=6$ min, the experimental and calculated values of Fd^2 compare as follows:

$$\text{From Fig. 3: } 101.5, 68.0, 58.8, 36.0, 31.7, 21.3.$$

$$\text{From Eq. (1): } 101.1, 67.8, 59.2, 36.2, 31.6, 21.2.$$

$$q_1 \epsilon^{-1} = 4.30 \text{ units,}$$

$$q_3 \epsilon^{-1} = 7.35 \text{ units,}$$

$$q_2 \epsilon^{-1} = 13.76 \text{ units,}$$

$$q_4 \epsilon^{-1} = 4.93 \text{ units.}$$

We see that Eq. (1) is confirmed within small limits of error. A more extensive experiment with n electrified particles would correlate $\frac{1}{2}n(n-1)$ values of Fd^2 in terms of only n values of q . We may note that an experiment with only three electrified particles would give us no information, one way or the other, concerning the validity of Eq. (1).

Having demonstrated the preceding correlation in terms of a single-valued quantity q for each particle, we admit q as an accepted physical quantity to be expressed in specific units and give it the name *electric charge*, or simply *charge*. Since

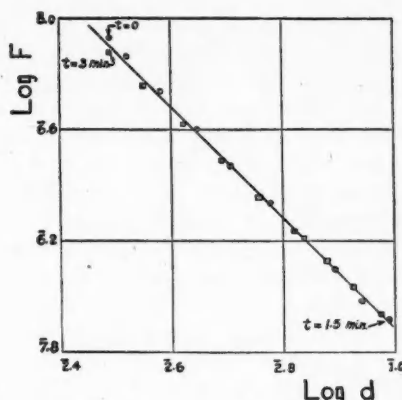


FIG. 2. Graph indicating an inverse-square law of force between two electrified pith balls. The distance d between centers is expressed in meters and the mutual force F in newtons. Points surrounded by circles were obtained from photographs taken from time $t=0$ to 1.5 min; and points surrounded by squares, from $t=1.5$ min to 3 min. When $t=1.5$ min, $Fd^2 = 76 \times 10^{-10}$ newton m^2 .

the basic demonstrations are performed in air, it would seem advisable at this point to anticipate the results of measurements on capacitors with different insulating materials, in which it is found that the proportionality factor has different relative values for different mediums and that $\epsilon_a/\epsilon_0 = 1$, to a very close approximation, where ϵ_a refers to air and ϵ_0 to vacuum; we call ϵ_a the *permittivity of air* and ϵ_0 the *permittivity of space*.²

The mks unit of charge, the coulomb, may be defined tentatively by setting $\epsilon_0 = 1/(9 \times 10^9)$ coul² newton⁻¹ m⁻² (this seemingly odd value may need some explanation). For the charge on pith ball No. 1 at time $t=2$ min, for example, we get $q_1 = 4.69 \times 10^{-10}$ coul. We note that charge is defined only for the particles of a "system" and that the system must contain more than three electrified particles for the definition to have any significance.

It is usual to explain at an early stage that electrons and nuclei are the ultimate sources of electric charge and that the Coulomb law, Eq. (1), applies in general only to elementary particles. Applicability of Eq. (1) to a pair of pith balls,

² At some time or other, it might be well to point out that: (i) the use of characteristic values of ϵ for material mediums is merely a convenient means of escape from (or, if one prefers, is equivalent to) summing electric force over all electrons and nuclei of the medium, for which summation we would need only the one proportionality factor ϵ_0 ; (ii) this practice has definite limitations; see, for example, C. C. Murdock, *Am. J. Physics* 12, 201 (1944).

where d is taken from center to center, may be explained as follows: (i) the electrified state of each pith ball is one in which there is a uniform, or very nearly uniform, distribution of electrons (or positively charged atoms, as the case may be) over its spherical surface; (ii) the force F in this instance is the resultant of all the individual forces exerted on the charged particles on the surface of one ball by the charged particles on the surface of the other ball, which force, for a uniform spherical distribution of charge, can be shown by integration to be the same as if the total charge on each ball were concentrated at its center. In any case, there is no thought of representing the pith ball experiments as a final basis on which the Coulomb law rests; they are presented primarily for their instructional value. The really compelling evidence for the validity of the Coulomb law is the complete self-consistency of electrostatic theory, founded on this law.³

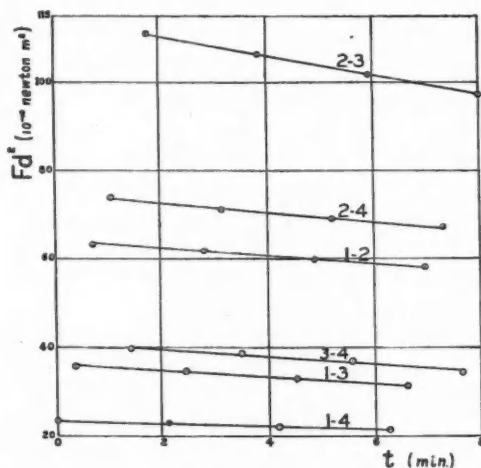


FIG. 3. Values of Fd^2 versus time t for successive interactions between four pith balls, Nos. 1, 2, 3 and 4, brought near each other in pairs.

³ Most of the evidence is indirect. For example, prior to the torsion-balance experiment of Coulomb (1785), Cavendish (1772-73) arrived at the inverse-square relation by ingenious deduction from an experiment with concentric spheres and a pith-ball electroscope, a predecessor of the Faraday ice-pail experiment. One of his manuscripts [see J. C. Maxwell, ed., *The electrical researches of the Honourable Henry Cavendish* (Cambridge, 1879), pp. 104-112] states: "We may therefore conclude that the electric attraction and repulsion must be inversely as some power of the distance between that of the $2+1/50$ th and that of the $2-1/50$ th, and there is no reason to think that it differs at all from the inverse duplicate ratio." Maxwell (pp. 417-19) performed a similar experiment with an improved arrange-

The Coulomb law can be expressed more fully in terms of vectors. We have

$$\mathbf{F}_{x-y} = q_x q_y \mathbf{r}_{x-y} / \epsilon_0 r_{x-y}^3, \quad (8)$$

where \mathbf{F}_{x-y} is the individual electric force exerted on any particle x by any other particle y , and \mathbf{r}_{x-y} is the position vector of particle x relative to particle y ; q_x and q_y may be either positive or negative. Equation (8) is an extension of Newton's third law, giving more detailed information for the particular case of an electric interaction (forming \mathbf{F}_{y-x} in similar fashion, we get $\mathbf{F}_{y-x} = -\mathbf{F}_{x-y}$). The Coulomb law, like other extended forms of Newton's third law such as the law of gravitation, has no meaning in a strict sense except in conjunction with the law of superposition. The set of complementary relations—namely, the law of superposition plus Newton's third law or any or all of its extended forms—which we apply in practice I have previously chosen to call the *Mach law of inertia*.⁴

In the Ampère law for the individual force $d\mathbf{F}_{x-y}$ exerted on a current element $i_x ds_x$ by another current element $i_y ds_y$,

$$d\mathbf{F}_{x-y} = \mu_0 i_x i_y ds_x \times (ds_y \times \mathbf{r}_{x-y}) / r_{x-y}^3,$$

we encounter another proportionality factor μ_0 , called the *permeability of space*. In the Maxwell field theory, μ_0 turns out to be related to ϵ_0 and c , the speed of radiation in space, by the equation $\epsilon_0 \mu_0 = 1/c^2$. It happens that μ_0 is more directly concerned than ϵ_0 in the measurements in which the absolute values of practical standards are determined.⁵ It is therefore preferable in the end to take μ_0 as the primary standard magnitude in place of ϵ_0 ; the mks value assigned to it is $\mu_0 = 10^{-7} \text{ kg m coul}^{-2} = 10^{-7} \text{ h m}^{-1}$, and so forth. Permittivity of space is then a secondary magnitude whose value depends on the values of μ_0 and c ; taking $c = 3 \times 10^8 \text{ m sec}^{-1}$, we get the value of ϵ_0 previously given.

Concerning Dimensions

There would seem to be room for discussion, beyond what has been written by others, con-

cerning the use of a Thomson quadrant electrometer. He says, "We may therefore conclude that g , the excess of the true index above 2, must either be zero, or must differ from zero by less than $\pm 1/21600$."

⁴ A. J. O'Leary, *Am. J. Physics* 15, 146, 336 (1947).

⁵ L. J. Briggs, "The national standards of measurement," *Rev. Mod. Physics* 11, 111 (1939).

cerning differences in the dimensions of quantities in the different systems of units in electrodynamics. The question may be clarified, I think, through a comparison of the definition of charge with the definition of a primary quantity in dynamics.

In the objective approach of Mach, mass is defined⁶ according to a method that is essentially operational, yielding the equation

$$\mathbf{a}_{x-y}/\mathbf{a}_{y-z} = -m_y/m_x \quad (9)$$

for an interaction between any two particles x and y . Unit mass is defined by assigning a particular mass value to the International Kilogram. The significance of mass is extended in the law of gravitation; for a gravitational interaction between two particles x and y , we have

$$\mathbf{a}_{x-y} = Gm_y\mathbf{r}_{y-x}/r_{y-x}^3 \quad (10)$$

and

$$\mathbf{a}_{y-x} = Gm_x\mathbf{r}_{x-y}/r_{x-y}^3, \quad (11)$$

where $\mathbf{r}_{x-y} = -\mathbf{r}_{y-x}$, and G is a proportionality factor known to be a constant.⁷ If we should choose to do so, we might dispense with a primary material standard and define unit mass by making G a primary standard magnitude; for example, we might define an absolute mass unit, the *abmu*, by setting $G = 10^{-11} \text{ abmu}^{-1} \text{ m}^3 \text{ sec}^{-2}$. In this alternative procedure, the definition of mass for the particles of a system *relative to one another* in accordance with Eqs. (9)–(11), and our recognition of a third physical dimension, would not be changed in any way. Except that it would not be practicable to do so, the International Kilogram might be relegated to the role of a secondary practical standard. The absolute value of its mass would be measured in a gravitational experiment of the Cavendish type; the measurements of Heil yield the value $6.673 \pm 0.003 \text{ abmu}$.

The preceding approach to the definition of charge is similar to the method of defining mass;

we find that we can correlate the values of Fd^2 for different pairs of electrified particles in terms of a primary quantity *charge* characterizing each particle.⁸ There are several possible ways of defining unit charge. We might, for example, designate the electron as a standard particle, adopt the charge q_e characteristic of an electron as a primary standard magnitude, and define unit charge by assigning a particular value to q_e . It happens to be more practicable to do without a primary material standard; we define electric units by assigning a particular value to either ϵ_0 or μ_0 . In the mks system of units, a fourth physical dimension is clearly recognized; μ_0 is made a primary standard magnitude, distinct from a numeric. On the other hand, in the cgs-es system of units, unit charge is defined by setting $\epsilon_0 = 1$ without dimensions; and in the cgs-em system, the various electric units are established by setting $\mu_0 = 1$ without dimensions. The lack of a fourth physical dimension in the latter two systems, it seems to me, is definitely inconsistent with our treatment of dimensions in mechanics; it is an oversight that should be rectified. A particular choice in the method of defining unit charge does not in any way change the definition of charge for the particles of a system *relative to one another* in accordance with the Coulomb law and should not affect our recognition of a new physical dimension. Since the relations between quantities are the same in all systems of units (except for certain numerical factors depending on whether a system is rationalized or not and certain minor differences in form arising from the value unity for ϵ_0 or μ_0), the relative dimensions of all quantities should be made the same in all systems; for example, $[\epsilon_0] = [L^{-2}T^2\mu_0^{-1}]$, $[\mu_0] = [L^{-2}T^2\epsilon_0^{-1}]$, $[q] = [L^{\frac{1}{2}}M^{\frac{1}{2}}\mu_0^{-\frac{1}{2}}]$, and so forth.

* * *

I wish to thank Mr. George Kayser for constructing the equipment and for his help in taking photographs and making measurements.

⁸ I am restricting attention to the customary development of electrodynamics from the Coulomb law as the starting point. It would be possible, though more difficult, to introduce electrodynamics (not merely a system of units) from some other experimental basis, say the Ampère law. Whatever starting point is used, a fourth primary quantity must be defined in terms of operations by which it is possible to measure it in units of its own kind, in distinction from the method of defining secondary quantities.

⁶ For full details, see reference 4.

⁷ The law of gravitation, with G initially taken to be a constant, is commonly said to define *gravitational mass* in distinction from the quantity *inertial mass* appearing in Eq. (9); in view of this distinction, experiments such as those of Eötvös [*Ann. d. Physik* 68, 11 (1922)] are said to indicate the equivalence of inertial mass and gravitational mass for all bodies. Another interpretation with a slightly different shade of meaning is this: taking Eqs. (10) and (11) initially to express a law concerning the same quantity mass as in Eq. (9), experiments such as those of Eötvös indicate that G is a constant, the same for all pairs of particles independently of their nature or state.

Reproductions of Prints, Drawings and Paintings of Interest in the History of Physics

33. *Quasi Cursores Vitai Lampada Tradunt*

E. C. WATSON

California Institute of Technology, Pasadena 4, California

IN 1884, in connection with the celebration of the tercentenary of the founding of the University of Edinburgh, there was issued a collection¹ of portraits of the officers and professors of the university drawn and etched by WILLIAM HOLE. Excellent full-page likenesses are given of 39 professors, and most of them are shown lecturing to students or at work in their offices or laboratories. Of these the five reproduced in this article are of especial interest to physicists.

The chemist LYON PLAYFAIR (Plate 1) is perhaps best known as the first Member of Parliament for the Universities of Edinburgh and St. Andrews. He was, however, a competent chemist and held the Chair of Chemistry at

Edinburgh from 1856 to 1869. As "Jehu Junior"² wrote in 1875 in connection with "Ape's"³ caricature of him in *Vanity Fair*,

He was born five-and-fifty years ago in India; he took early to Scotland and chemistry, he engaged naturally in calico-printing, and he became a chemical professor in Manchester. In his particular vocation he grew to be an authority, was set to do various things by the Government, and, being adopted as a Special Commissioner in the Exhibition of 1851, he blossomed into a Companion of the Bath and a Court functionary, so far as the Prince Consort could make him one. Of course he became one of the Department of Science and Art when that was invented, and he has been made a member of many learned societies, a Knight of many Orders, and a Doctor of Laws by the University of Edinburgh. He believes truly in nothing but chemistry, but in that he believes always and everywhere. He has attached himself to Ozokerit candles and Liberalism, and thus has become a Member of Parliament, where he has been believed to be distinguished without ever having proved that he is a statesman. Wherefore he was made a Postmaster-General, for which he proved his peculiar fitness by discovering that if penny stamps had been blue instead of red he could have saved the country thirty thousand a year. Nevertheless, before he took office he was wont to attempt criticism of the Liberal Leader; but although he always assumed to exercise an independent judgment, it always led him to give a dependent vote. He is an excellent Professor.

The chemist ALEXANDER CRUM BROWN (Plate 2) did not attain sufficient fame to be listed in the *Dictionary of National Biography*; but he was a theorist well acquainted with the history and theories of chemistry of his day, and he apparently just missed inclusion along with KEKULE and KOLBE in the first rank of theoretical chemists of the period. He was a Fellow of the Royal Society and one of the Vice Presidents of the Chemical Society, but he published very little and had no interest in analytical or practical chemistry. He did not, however, confine himself to chemistry, and his knowledge of mathematics,



PLATE 1. LYON PLAYFAIR (1818–1898).

¹ *Quasi cursores. Portraits of the high officials and professors of the University of Edinburgh at its tercentenary festival* (University Press, Edinburgh, 1884).

² Thomas Gibson Bowles.

³ Carlo Pellegrini.



PLATE 2. ALEXANDER CRUM BROWN.

philology, modern languages, even such as Russian and Chinese, and of church history was well known to his friends.

The engineer FLEEMING JENKIN (Plate 3) is best known for his work with KELVIN on submarine cables and with MAXWELL on electrical standards, but he also established an excellent School of Engineering at Edinburgh and at the same time did quite varied scientific and literary work.

Very plain-featured, rather short in stature, always youthful and energetic in manner, Jenkin did not prepossess strangers, and his flow of words and love of disputation never made him very popular. As a lecturer he was interesting, and he was a good disciplinarian. His taste in literature was broad and unconventional, and he exhibited a sound critical faculty in his miscellaneous essays and reviews. He was also an excellent amateur actor and dramatic critic. In practical engineering, thoroughness and soundness marked all his work. His determinative work in electricity is of the highest value, while his varied originality as an inventor is testified by his thirty-five British patents.⁴

⁴ G. T. Bettany in the *Dictionary of national biography*.

ROBERT LOUIS STEVENSON wrote the memoir⁵ that was prefixed to JENKIN's collected papers⁶ and described his conversation in the well-known essay on "Talk and Talkers" as follows:

His manner is dry, brisk and pertinacious, and the choice of words not much. The point about him is his extraordinary readiness and spirit. You can propound nothing but he has either a theory about it ready-made, or will have one instantly on the stocks, and proceed to lay its timbers and launch it in your presence. "Let me see," he will say. "Give me a moment. I *should* have some theory for that." A blither spectacle than the vigour with which he sets about the task, it were hard to fancy. He is possessed by a demoniac energy, welding the elements for his life, and bending ideas, as an athlete bends a horseshoe, with a visible and lively effort. He has, in theorising, a compass, an art; what I would call the synthetic gusto; something of a Herbert Spencer, who should see the fun of the thing. You are not bound, and no more is he, to place your faith in these brand-new opinions. But some of them are right enough, durable even for life; and the poorest serve for a cock-shy—as



PLATE 3. FLEEMING JENKIN (1833-1885).

⁵ The memoir of *Fleeming Jenkin* will be found in almost any set of Stevenson's collected works.

⁶ Two volumes (London, 1887).



PLATE 4. GEORGE CHRYSTAL (1851–1911).

when idle people, after picnics, float a bottle on a pond and have an hour's diversion ere it sinks. Whichever they are, serious opinions or humours of the moment, he still defends his ventures with indefatigable wit and spirit, hitting savagely himself, but taking punishment like a man. He knows and never forgets that people talk, first of all, for the sake of talking; conducts himself in the ring, to use the old slang, like a thorough "glutton," and honestly enjoys a telling facer from his adversary. Cockshot [*i.e.*, Jenkin] is bottled effervescency, the sworn foe of sleep. Three-in-the-morning Cockshot, says a victim. His talk is like the driest of all imaginable dry champagnes. Sleight of hand and inimitable quickness are the qualities by which he lives.

The mathematician GEORGE CHRYSTAL (Plate 4), after graduating from Cambridge as second wrangler and Smith's Prizeman, applied himself to the study of physics under MAXWELL in the Cavendish Laboratory and wrote the articles on "Electricity" and "Magnetism" for the ninth edition of the *Encyclopaedia Britannica*. After his election to the Chair of Mathematics at Edinburgh he still continued his experimental researches and published a number of valuable

experimental, as well as mathematical, papers. He was noted for the clearness and conciseness of his expositions. His textbook on *Algebra* went through many editions and is still a standard work.

Better known to physicists than any of these, however, is PETER GUTHRIE TAIT (Plate 5), co-author with KELVIN of the great *Treatise on Natural Philosophy* and one of the most prolific physicists of the Victorian Era.⁷ He was the personal friend of HAMILTON, ANDREWS, STOKES, JOULE, KELVIN, MAXWELL, HELMHOLTZ, CAYLEY and many others of the great galaxy of Victorian physicists and mathematicians. His correspondence, much of which fortunately has been preserved,⁸ throws a bright and very pleasing light upon the personalities of this unusual group of men.

Towards Hamilton Tait was the loyal disciple, eager to have the master's help at all stages, and always



PLATE 5. PETER GUTHRIE TAIT (1831–1901).

⁷ He wrote at least 365 papers, both theoretical and experimental, and 22 books.

⁸ See C. G. Knott, *Life and scientific work of Peter Guthrie Tait* (Cambridge, 1911).

ready to give him the fullest credit as the prime source of every luminous thought. . . . To Cayley Tait turned as to the embodiment of mathematical wisdom and knowledge. . . . But for Maxwell Tait had not only unstinted admiration as a man of science; he had for him a deep strong love which had its roots in common school life and grew, strengthened and ripened with the years. He understood to the full Maxwell's intellectual oddities, his peculiar playful humour, his nobility of character, and the deeper thoughts which moved his mind but rarely found expression.⁹

As a lecturer TAIT was probably unsurpassed by any of his contemporaries. His tall figure and magnificent head at once impressed his audiences, and the impression was deepened by the "easy utterance, clear enunciation, and incisive phrase" that characterized his expositions. It was probably this quality which won for him over MAXWELL the appointment at Edinburgh. As KNOTT says,¹⁰ "Maxwell towers as one of the

creative geniuses of all time, curiously lacking though he was in the power of oral exposition; Tait, who possessed, also by intuition, the clearest physical conceptions, has left behind him a great record of research both in mathematics and physics, while, as a teacher and clear exponent of physical laws and principles, he took a foremost place among his contemporaries." It is appropriate therefore that the artist chose to depict him at his lecture table. Unfortunately, as KNOTT states,¹¹ HOLE, "although very happy in most of his delineations [in *Quasi Cursorae*], has not caught Tait quite satisfactorily. The attitude and figure generally are admirable, as are also the accessories of the Holtz machine, Leyden jars, and blackboard; but the expression of the face is not altogether suggestive to those who knew him well."

⁹ C. G. Knott, reference 8, p. 166.

¹⁰ Reference 8, p. 17.

¹¹ Reference 8, pp. 49-50.

He [Sir Humphry Davy] commonly wrote his lectures the day before he delivered them. He was always master of his subject, and composed with great rapidity and with a security of his powers never failing him. Latterly he trusted a good deal to notes, and, excepting on particular occasions, wrote little more than the parts which he wished to make most impressive. It was almost an invariable rule with him, the evening before, to rehearse his lecture in the presence of his assistants, the preparations having been made and everything in readiness for the experiments; and this he did, not only with a view to the success of the experiments and the dexterity of his assistants, but also in regard to his own discourse, the effect of which, he knew, depended on the manner in which it was delivered. He used, I remember, at this recital, to mark the words which required emphasis, and study the effect of intonation, often repeating a passage two or three different times to witness the difference of effect of variations in the voice. His manner was perfectly natural, animated, and energetic, but not in the least theatrical—he spoke as if devoted to his subject. . . . His experiments were devised on the same principle, not of amusing and pleasing, but of illustrating his discourse, and demonstrating either important properties of bodies or principles of the chemical action of bodies; he took every care that they should not appear to have been introduced for show and to excite merely wonder, even when most brilliant and wonderful.—JOHN DAVY, *Memoirs of Sir Humphry Davy* (1839), vol. 1, p. 92.

NOTES AND DISCUSSION

A Graphical Treatment of the Physical Pendulum Problem

ERIC J. IRONS

Queen Mary College, University of London, London, England

LET Fig. 1 represent a physical pendulum whose center of gravity and point of suspension are, respectively at G and at O_1 , and let $O_1G = h$. By the usual elementary theory it is readily shown that the period of the pendulum is given by the equation

$$T = 2\pi(I/Mgh)^{1/2} = 2\pi(K^2/gh)^{1/2} = 2\pi[(h^2 + k^2)/gh]^{1/2} = 2\pi(l/g)^{1/2},$$

where K and k are, respectively, the radii of gyration about O_1 and G , and l is the length of the equivalent ideal pendulum. It is the purpose of this note to describe a graphical method of tracing the variation of l (and, incidentally, of T) with h .

To this end describe a circle having center G and radius k , and draw $GN[=k]$ perpendicular to GO_1 . Then $O_1N = (h^2 + k^2)^{1/2} = K$. Next join P , the point of intersection of O_1N and the circle, to G and produce PG to cut the circle again in Q ; join also Q and N and let the resulting line cut O_1G produced in O_1' . Then, as angle O_1NO_1' is a right angle,

$$\cos GO_1N = h/K = K/O_1O_1',$$

or $O_1O_1' = K^2/h = l$. Thus O_1' is the center of oscillation; and, if the body were suspended from it, the corresponding periodic time would be

$$T' = 2\pi(K'^2/gh')^{1/2} = 2\pi(O_1O_1'/g)^{1/2} = T,$$

since $\cos GO_1'N = h'/K' = K'/O_1O_1'$.

It follows that $l = O_1O_1'$ when $h = O_1G$ (or $h' = O_1'G$), and that two points may be located on the curve of l versus h (or h') by erecting ordinates of length O_1O_1' at O_1 and O_1' . Other points corresponding to O_2 , O_2' ; O_3 , O_3' ; . . . may be determined similarly; and, as the curves are symmetrical about NG produced, they may be completed.

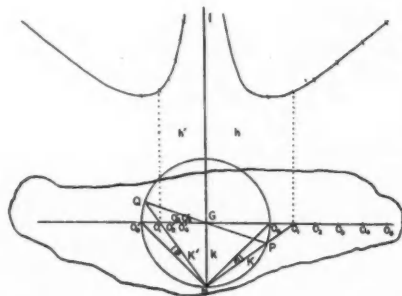


FIG. 1. Physical pendulum. Note: The l at the upper right-hand side of the figure should read l , the k immediately to the right of the vertical center line is read h .

Assuming the value of g , a curve of T^2 versus h is readily obtained by multiplying the scale of l by the appropriate factor; or, alternatively, some values of T may be marked on the l axis.

Further, the diagram shows that $O_1O_1'/\sin\theta = K/\sin 135^\circ$; $O_1O_1'/\sin\phi = K'/\sin 45^\circ$; $\theta = \phi$ (as there are two right angles at N); and $K > K'$. Hence $O_1O_1' > O_1'O_1'$, or $O_1O_1' > O_1'O_1'$; that is, l has a minimum value, namely $2k$, when $h = h' = k$.

Koenig's Interference Apparatus

PAUL F. GAHR

Wells College, Aurora, New York

IN an interesting article¹ on manometric flames which recently appeared in this Journal, mention was made of the use of manometric flames in Koenig's interference apparatus. The authors of that paper stated that, using rotating mirrors and manometric flames, the sliding tube could be set for interference to a certainty of 1 percent. The apparatus, when thus used, is not suitable for lecture demonstration because the flames are hard to see, and because precision requires too much time. It seemed to us that the apparatus is well worth showing provided that everything is visible, and that measurements can be quickly made with adequate precision. In order to show both constructive and destructive interference, for several different values of path difference, we used, as source, a high pitched whistle;² the intensity of the resultant sound waves was indicated on the lecture-room galvanometer. Whereas the entire process is described below, only the demonstration of interference was shown, and the calculation of the velocity of sound.

To the Y-piece at the exit end of the interference apparatus, in which are re-united the sound waves transmitted by the two paths, was attached a telephone transmitter. The transmitter was in series with a battery and the primary of a telephone induction coil. The secondary of this coil could be connected either to the vertical deflectors of an oscilloscope or, through an instrument rectifier, to the galvanometer. The galvanometer was arranged to be dead-beat, and to give a deflection near the end of the scale for a sound maximum. The horizontal deflectors of the oscilloscope were connected to the time sweep and were synchronized to a microphone hummer, whose frequency had been accurately determined to be 1052.5 cy/sec. The whistle was blown from the compressed-air supply, a water manometer giving the pressure at the whistle. We found by trial a pressure which gave a good tone and slight variations of which affected the pitch unappreciably. With the output connected to the oscilloscope, the length of the

whistle was adjusted until three waves were obtained on the oscilloscope. Hence the frequency of the whistle was known to be 3157.5 cy/sec at the chosen air pressure.

With the output changed to the galvanometer, the slide on the interference apparatus was drawn out until the galvanometer indicated a maximum, and the scale reading of the slide was carefully noted. Similarly the adjacent position for a minimum was noted, and so on. Altogether six maxima and five minima could be obtained, forming a very convincing arithmetic series. The writer made a demonstration of the maxima in 6 min, although he had to watch the air pressure and the galvanometer. The following set of readings was obtained for the maxima: 6.7, 12.0, 17.5, 23.0, 28.3, 33.9 cm. Remembering that changes in scale reading must be multiplied by two to find increases in path length, we computed² $\lambda = 10.89$; hence the velocity of sound at 24°C is 344.87 m/sec, and the velocity at 0°C is 330.23 m/sec. The tube settings for a maximum were accurate to within 1 mm, and the settings for a minimum were somewhat better. The accuracy of the final result is better than 1 percent.

¹ R. J. Stephenson and G. K. Schoepfle, *Am. J. Physics* 14, 294 (1946).

² R. M. Sutton, *Experiments in physics* (McGraw-Hill, 1938), p. 158.

³ P. F. Gaehr, "Equations for straight lines," *Am. J. Physics* 15, 430 (1947).

A Mousetrap Atomic Bomb

RICHARD M. SUTTON

Haverford College, Haverford, Pennsylvania

MORE people are familiar with mousetraps than with atoms. And most people would suppose that any resemblance between mousetraps and atomic bombs is purely accidental. It is. But even accidental resemblances are often interesting and instructive. Professor L. E. Dodd has described a concatenation of mousetraps,¹ and the present note presents a modification and extension of his idea.

When a mousetrap is set, it stores about 8×10^6 ergs of potential energy ready for prompt delivery when its trigger is touched. An atom of uranium-235 stores about 3×10^{-4} ergs of releasable potential energy in its nuclear structure, subject to release by the chance entry of a neutron. Atoms of uranium and plutonium, though relatively stable, may be triggered by neutrons. Fission results with release of energy and ejection of more neutrons. If fissionable atoms in sufficient numbers are present, and if too many neutrons are not lost from the neighborhood, a chain reaction may result with one atom setting off some one or more of its neighbors, and these in turn setting off others. Chain reactions are *chance* reactions of high probability; they are not linked reactions with a one-to-one relationship quite as simple and definite as the links of a chain.

In the mechanical analog described here, the "atoms" are mousetraps and the "neutrons" are cork stoppers. Potential energy is stored in each mousetrap-nucleus, by

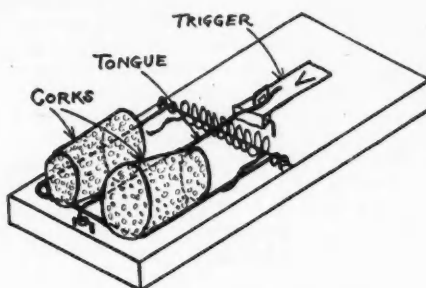


FIG. 1. Mousetrap-"atom" with "neutrons."

setting the spring and trigger, and two or three cork stoppers are then placed on each set spring on either side of the tongue (Fig. 1). If the trigger of any trap is touched, the trap snaps violently and throws the neutron-corks placed upon it.

It remains only to provide enough traps in close proximity to make the explosion of one trap *likely* to affect some nearby trap. In this case, the chain reaction is based purely on chance and there is no special causal relationship between one trap and any of its neighbors. But if part of the lecture table is covered with three or four dozen traps ready for action in the way described, and if a cork is tossed into their midst, violent and sudden action usually follows. After the explosion ends, many traps will be found undisturbed, corresponding to uranium atoms that have escaped the cascade. A chain of higher probability may be assured by placing a sheet of cardboard horizontally over the table about 1 ft above the traps.² Then, those neutron-corks which would otherwise be shot across the room will be reflected to the table, and the chance of their springing other mousetrap-atoms is greatly increased as the "flux of neutrons" is made greater. It may be objected that there are no good reflectors of neutrons in nature, but remember that this is just a single-layer bomb and that the reflecting sheet may be considered in its effect as equivalent to other virtual layers of atoms situated above, each contributing to the reaction. In fact, the relation of partial explosion of the traps in the absence of the reflector to almost complete explosion with the reflector present is closely linked with the concept of "critical size" in the atomic case. More traps, more chance of continued action. At best, the number of traps is an insignificant fraction of the Avogadro number.

If a still better mechanical representation of nuclear fission is desired, simply double the number of traps! Place one trap upon another so that the upper trap as a whole will be flung when the lower trap is sprung. So violent is the motion of a trap that only in rare cases will either member of a pair spring without springing the other as well. Fission is then nearly complete: mousetraps and corks fly in all directions. If more pains are taken, the traps may be set in layers on a "space lattice" with the layers 6 or 8 in. apart. In this case, the trigger of each trap should have an appendage that projects through a hole bored in the wooden base. It can then be sprung by im-

pact from above or below. Cork-neutrons shot upward may spring traps above them, as well as those on the same or a lower level, provided that the traps are placed on the lattice to allow passage of corks through gaps. About six layers of two dozen traps each should make a gratifying display of the random but wholesale sequential release of potential energy.

The space-lattice arrangement of mousetrap-atoms also suggests an analogy with the nuclear pile in which fissionable materials are prepared and atomic energy is released under controlled conditions. If cardboard barriers are inserted between successive layers of traps, then the chance explosion or partial explosion of one layer will not affect adjacent layers. The barrier sheets serve to stop (but not absorb) the neutron-corks in much the manner that boron rods regulate the reaction in the uranium-carbon pile by absorbing neutrons.

Finally, anyone who is not satisfied with mousetraps and corks may use rat-traps and baseballs!

¹ In an unpublished paper that was available to me in manuscript when I presented several of his experiments at the annual meeting of the Association in New York, in January 1946.

² Explosion of 60-70 percent of the traps is usual in this case.

On the Use of the Term "Voltage"

RALPH G. HUDSON

Department of Electrical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts

THERE are several basic reasons for replying in the negative to Doctor Hazeltine's question, "Is not voltage a desirable term?"¹ and for supporting Doctor Roller's disparagement of the term.²

Since the Act of Congress of July 12, 1894 definitely specifies "The unit of electromotive force" and makes no mention of "voltage," it is doubtful if the latter term has any legal standing. In a case before the Supreme Court regarding the contested rating of a 115-v, 60-w lamp, the definition of 115 v must be based upon the applied "electromotive force" and not the "voltage."

The term "voltage" is moreover inconsistent. Why not "ohmage," "amperage," "kilowatthourage," "microfaradage," "henryage," "gaussage," "ampere turnage," and so forth? To be sure, some of these terms are used by addicts of "voltage," but with no more justification.

The single word "voltage" substituted for the widely different meanings of "emf" and "potential" is furthermore ambiguous and confusing. To quote from the *Standard handbook for electrical engineers*, "Emf is ordinarily accompanied by a difference of electric potential; but an emf may occur without difference of potential, as for example, when a straight bar magnet is thrust symmetrically into a circular loop of uniform wire. A brief current will thus be set up in the wire due to an emf induced by the magnet's motion; but there will be no difference of electric potential in or around the wire."

In the foregoing illustration the adherent of "voltage" must explain the situation by stating that one kind of

voltage is present but the other kind of voltage is not, which is no explanation at all. Practically every question concerning electrical apparatus is veiled with uncertainty if "voltage" is used. For example, "What is the voltage of a transformer or a storage battery?" Any answer could be interpreted either as "emf" or as the "potential" between the terminals.

Some sponsors of "voltage" excuse the adoption by maintaining that "electromotive force" is too long. This contention is invalidated by the fact that "emf" is practically always used and is a shorter word than "voltage." The use of "voltage" also requires more space or time because it must usually be preceded by an explanatory adjective if it is to be given an intelligible meaning.

¹ A. Hazeltine, *Am. J. Physics* 15, 191 (1947).

² D. Roller, *Am. J. Physics* 14, 340 (1946).

Determining the Positions of Maximum Intensity in the Single-Slit Diffraction Pattern

DONALD V. BARRER

Ripon College, Ripon, Wisconsin

THE spacing of the bands of maximum intensity in the Fraunhofer single-slit diffraction pattern is shown in many textbooks¹ to be proportional to the roots of the equation $\tan x = x$. Frequently a number of these roots are tabulated, and credit for their computation is given to Schwerdt.² This is misleading since reference to his work reveals that he estimated the positions of only the first two maxima from a graph of the intensity function. Occasionally the reader is referred to Table I, page 145

TABLE I. Roots of $\tan x = x$.

n	Textbook value	Value by present method*
1	1.4303 π	1.4302969 π
2	2.4590 π	2.4590242 π
3	3.4709 π	3.4708898 π
4	4.4774 π	4.4774177 π
5	5.4818 π	5.4815369 π
6	6.4844 π	6.4843843 π
7	7.4865 π	7.4864750 π
8		8.4880689 π
9		9.4893268 π
10		10.4903436 π
100		100.499032 π

* Calculations were made using Bruhns, *A new manual of logarithms* (Chas. T. Pownier Co., Chicago, rev. ed., 1941).

in Schwerdt's book, which is a tabulation of the relative intensity for five-degree increments rather than a table of positions of maximum intensity.

Solutions of the equation $\tan x = x$ can be effected by Newton's method. A rough graph shows that there are roots lying near $3\pi/2$, $5\pi/2$, $7\pi/2$, \dots , and these values may be used as first approximations if the equation is written in the form $\sin x = x \cos x$. If accuracy beyond one or two decimal places is desired the computation involved is rather tedious. To compute the roots shown in Table I of this note the following somewhat different method was used, which lessened the work considerably.

In the equation

$$\tan x = x \quad (1)$$

set

$$x = 3\pi/2 - y, \quad (2)$$

obtaining

$$\tan(3\pi/2 - y) = 3\pi/2 - y$$

or

$$\cot y = 3\pi/2 - y. \quad (3)$$

Since y is small, make the approximation, $\cot y \approx 1/y$, thus obtaining the quadratic equation

$$y^2 - (3\pi/2)y + 1 \approx 0. \quad (4)$$

Then use the root

$$y \approx \frac{1}{2}[3\pi - (9\pi^2 - 16)^{1/2}] \quad (5)$$

as an approximation to y . Substituting from Eq. (5) in Eq. (2), we obtain the first approximation to the root,

$$x_1 = \frac{1}{2}[3\pi + (9\pi^2 - 16)^{1/2}]. \quad (6)$$

In a similar manner set $x = x_1 - y'$ and substitute in Eq. (1), expand $\tan(x_1 - y')$, replace $\tan y'$ by y' , and solve the resulting quadratic to obtain a second approximation to the root. The other roots may be calculated in a similar manner by replacing $3\pi/2$ in Eq. (2) by $5\pi/2$, $7\pi/2$, $9\pi/2$, It may be easily shown that the difference between the actual root and the second approximation so obtained is less than 4×10^{-8} rad.

In Table I are shown several roots calculated by this method, and compared with values listed in the textbooks. It is interesting to note that the roots corresponding to $n=5$ do not agree. I believe that the textbooks are in error here.

¹ For example, Preston, *The theory of light* (Macmillan, 1928); Taylor, *College manual of optics* (Ginn, 1924); Monk, *Light principles and experiments* (McGraw-Hill, 1937).

² Schwerdt, *Beugungerscheinungen* (Mannheim, 1835).

Procedures for Nominating and Electing Society Officers

JOSEPH D. ELDER

Wabash College, Crawfordsville, Indiana

AT the meeting of the Executive Committee of the Association in January 1947 the writer was requested to review the procedures of various other societies with regard to nominations and elections of officers. This action was precipitated by two factors. One was the fact that heretofore nominations have been requested of the membership of the Association about October 1 each year, and the election ballot has been sent out on December 1; there has therefore been too little time for the nominating committee to review the suggestions made by the members and make its nominations.¹ The second factor has been the fear, expressed by several members in the past, that there would presently be refusals of nominations for president if two candidates continue to be nominated for that office.

Accordingly, the writer addressed a series of questions to the Secretary of each of the member societies of the

American Institute of Physics and of the Mathematical Association of America. This note summarizes their answers; their cooperation is gratefully acknowledged.

The terms of office of president and vice president are one year in the American Physical Society, the Acoustical Society of America and the Mathematical Association; two years in the Optical Society of America and the Society of Rheology. The vice president always (though not automatically) succeeds to the presidency in the Physical Society, usually in the Optical Society and the Society of Rheology, and sometimes in the Mathematical Association. The Acoustical Society (which adopted new by-laws in May 1947) chooses as one of its officers a president-elect, who serves as assistant to the president and automatically succeeds at the end of the term.

In the Society of Rheology, with a relatively small membership, about 75 percent of the members exercise their privilege of voting in an election; in the Physical and Optical Societies some 50 percent vote; in the Mathematical Association, 20 to 25 percent. In recent elections of the Association of Physics Teachers the figure has been about 45 percent.

In all societies, space is provided on the election ballot so that members may write in the name of a person not nominated for an office, but no person has been elected by this method.² The Optical Society, when there are, say, two candidates for an office, prints the ballots in such a way that each name appears first on half of the ballots.

Candidates for the presidency are usually known by name to most members of each society, but this is not always true of candidates for other offices. The secretary of one society feels that there is danger in the practice of nominating only the best-known men for the presidency, for such men are almost sure of election but usually have too little time to do justice to the office.

A member of the Association has recommended placing on the ballot as a candidate for the presidency only the name of the out-going vice president. Unless the vice president succeeds to the higher office the Association loses the benefit of the background of experience with Association affairs that he acquires during his year of service.

Another member would like to see the same end accomplished by having the vice president automatically succeed to the presidency at the end of a one-year term; he wonders what point there is in asking members to select a president by ballot if they have agreed that there shall be only one candidate.

A third suggestion is that members could vote more intelligently if a brief biography of each candidate were printed, preferably on the ballot. It might be possible, for instance, to reproduce, without any change, the biographies appearing in *American Men of Science*. The space required for such biographies would not be so large as greatly to increase the expense of printing the ballots.

The Executive Committee is anxious to have further suggestions regarding procedure in nominating candidates and electing officers for the Association. The following questions arise from the present inquiry. (i) Should only one candidate be nominated for the presidency of the Association? (ii) Should the vice president automatically become a

candidate for the presidency? (iii) Should the vice president automatically succeed to the presidency? (iv) Should the terms of office of president and vice president be one year, as at present? (v) Would there be an advantage in printing on the ballot brief biographies of the candidates? (vi) Is it worth while to provide for the writing in of names on the election ballot? Members who wish to answer any of the foregoing questions, or to make suggestions or comments, either for improvement in the present procedure or expressing approval of it, are urged to send them—a postal card will serve—to the writer, who will compile them and present them to the Executive Committee.

¹At the 1934 annual meeting the nominating committee was instructed by the Executive Committee to use the nominating ballots of the members as a guide but not as a mandate.

²The secretary of one society hopes that write-ins will be abolished by constitutional amendment, because they never come near affecting the outcome of an election and are very costly in time of the tellers.

Equations for Straight Lines

PAUL F. GAHR

Wells College, Aurora, New York

EXAMINATION of numerous laboratory manuals, both elementary and advanced, indicates that the finding of the constants in the equation of the straight line from a graph by the method of least squares is no longer taught at either of these levels. This is undoubtedly due to the great labor involved and the consequent chance of arithmetic errors. Furthermore, it sometimes happens that a line drawn according to the method of least squares does not fit the points as well as does a line drawn by eye. A careful review of the theory of least squares shows that the deviations, the sum of the squares of which should be a minimum, should be measured normally to the line; actually we measure the deviations vertically, and assume that the abscissas are free from experimental error.

Usually a shorter and easier method is substituted for the calculation of the slope, but one seldom sees a logical basis for this method. Such a basis does exist if one set of coordinates is free from experimental error and forms an arithmetic progression. As an example, let us determine the period of a pendulum by noting the times of the 1st, 11th, 21st, ..., 71st transits (to the right) past the midpoint. If the ordinal numbers of the transits are plotted as abscissas, and the times as ordinates, the result should be a straight line, whose slope is equal to the period of the pendulum. Obviously, $t_{71} - t_{31} = T(71 - 31)$, and so on. The left-hand members of this and the other three similar equations can be averaged, the result being the value of $40T$, from which T can be found. More conveniently, the sum of the left-hand members is found by subtracting the sum of the first four t 's from the sum of the second four. This sum is divided by 4 to get the average, and then by 40. The denominator, 160, can also be quickly arrived at by subtracting the sum of the first four ordinal numbers from the sum of the second four; that is, $224 - 64 = 160$.

In the sum of the left-hand members, accidental errors in the group of the first four t 's are apt to cancel out among

themselves, and similarly in the group of the second four. Systematic errors in the one group cancel against those in the other group. Thus the remaining error has a good chance of being small. If more observations are taken, the numerator of the expression for T will be larger, being proportional to the square of the number of observations, and the remaining error will probably be smaller; in fact, the relative error is probably much smaller. Under the conditions specified, this method is rigorously correct.

However, even if both sets of coordinates are subject to experimental error and neither set forms an arithmetic progression, the foregoing method yields results as satisfactory as those of the method of least squares. Its justification is in the good results. In addition to finding the slope m , we can easily find the coordinates of the mean point $P_m[(\Sigma x)/n, (\Sigma y)/n]$, which is a point on the least-squares line; and so it is possible to compute the intercepts. We will illustrate by a set of data that is not particularly good.

x	y	
1.2	4.0	$m = 32.1/15.6 = 2.0577$;
2.0	6.0	$(x_0 + 4.525)/11.088 = 1/m$;
3.1	8.1	$x_0 = 0.864$;
4.0	10.2	$y = 2.0577(x + 0.864)$
		$= 2.0577x + 1.7766$.
4.9	12.2	
6.0	14.0	
6.9	16.1	
8.1	18.1	
25.9	60.4	Lower halves
10.3	28.3	Upper halves
15.6	32.1	Differences
8)36.2	88.7	Total sums
$x_m = 4.525$	$y_m = 11.088$	Mean point P_m

The least-squares equation is $y = 2.0463x + 2.013$. The ordinates at $x = 0, 5, 10$ for the two equations differ by 0.1, 0.2, 0.2, respectively.

The writer has made comparisons between the two methods on at least 100 sets of data obtained by elementary students and has always found the deviations between them to be well within the experimental error.

Elementary Theory of the Doppler Effects

ELLIOT T. BENEDIKT*

Rensselaer Polytechnic Institute, Troy, New York

THE explanation of the Doppler effects¹ given in textbooks is usually based on a qualitative discussion of the relative motion of waves and points. In the present paper, an analytic derivation of the formulas for the two most important cases is given, which have some advantages over the usual procedure. For the sake of simplicity, the

discussion will be limited to propagation of waves in one dimension; however, the methods employed can be easily generalized to the case of propagation in two or three dimensions, as well as to the Doppler effects not considered here.

A periodic phenomenon occurring at a certain point is said to be "propagated" if it is reproduced at a different point (possibly with reduced intensity, as an effect of the distance and absorption by the intervening medium) with a delay l/c , where l is the distance through which the periodic phenomenon is transmitted; c is defined as the phase velocity. If the phenomenon is simply periodic with frequency ν and is propagated along a given direction in a homogeneous medium, the propagated quantity y (displacement, pressure, electric field, strength, or some other quantity) will be given, for particular values of l and time t , by the equation

$$y = A \sin 2\pi \nu [t - (l/c) + \tau], \quad (1)$$

where A is the amplitude of y (which might be constant or a function of l), and τ is a constant. The wavelength λ is, by definition, the shortest distance between two points for which the phase of y is the same, and is therefore given by the formula

$$\lambda = c/\nu. \quad (2)$$

We shall define as "receiver" any system of points² that is subjected to the phenomenon under consideration. Let us introduce a system of coordinates x' rigidly connected with the receiver, and let x_0 be the distance between the source and the origin of this system of coordinates. If the receiver is at rest, the propagated phenomenon will be represented with respect to the latter by Eq. (1), with the substitution

$$l = x_0 + x'. \quad (3)$$

The resulting equation will represent a sinoidal (monochromatic) wave.

If the receiver is in motion with uniform velocity v (considered positive or negative according to whether the motion is away from or towards the source), Eq. (3) will still hold; however, x_0 will vary according to the law $x_0 = x'_0 + vt$, and therefore $l = x'_0 + vt + x'$. Substituting this value in Eq. (1), we obtain the equation

$$y = A \sin 2\pi \nu \{t - [(x' + vt + x'_0)/c] + \tau\},$$

which can be written in the form

$$y = A \sin 2\pi (1 - v/c) \nu \{t - [x'/(c-v)] + \tau'\},$$

where $\tau' = \tau - x'_0/(c-v)$. The latter equation indicates that, with respect to the receiver, the propagated phenomenon is periodic with frequency $(1 - v/c)\nu$, and has the phase velocity $c-v$. Applying Eq. (2), we see that $\lambda = (c-v)/(1 - v/c)\nu$, where λ is the wave length that would be obtained if the receiver were at rest.

Consider now the case in which the source is in motion with constant velocity u , considered positive or negative according to whether the source is moving away from or

towards the receiver. We shall use a system of coordinates x rigidly connected with the receiver; the distance between the source and the origin of this system will be given by $x_0 + ut$. The distance l through which the periodic phenomenon is being transmitted will now be equal to the distance between source and a given point of the receiver at the time t' of the emission of the disturbance which occurs at the given point at the time t . The time t' must satisfy the equation

$$(x + x_0 + ut')/c = t - t', \quad (4)$$

expressing the fact that during the interval of time $t - t'$ the disturbance has to cover the distance $x + x_0 + ut'$ between the source and the considered point of the receiver. From Eq. (4) we get $t' = (ct - x - x_0)/(c + u)$, and consequently

$$l = x + x_0 + ut' = (x + x_0 + ut)c/(c + u).$$

Substituting this value in Eq. (1) we obtain

$$\begin{aligned} y &= A \sin 2\pi \nu \{t - [(x + x_0 + ut)/(c + u)] + \tau\} \\ &= A \sin 2\pi \nu \{[t/(1 + u/c)] - [x/(c + u)] - [x_0/(c + u)] + \tau\} \\ &= A \sin 2\pi [(1 + u/c)\nu] \{t - (x/c) + \tau'\}, \end{aligned}$$

where $\tau' = (1 + u/c)\tau - (x_0/c)$. This equation for y shows that the propagated phenomenon has, with respect to the receiver, a frequency $\nu/(1 + u/c)$ and is propagated with a phase velocity³ c . Applying Eq. (2), we find the wavelength to be $(1 + u/c)c/\nu = (1 + u/c)\lambda$, where λ is the wavelength that would be obtained if the source were at rest.

As for the range of validity of these results, it should be noted that if v or u become much larger than $\geq c$, the foregoing formulas lose physical meaning. In the case of a moving receiver, we would obtain a negative speed of propagation of the disturbance with respect to the receiver, and a negative frequency if the motion is away from the receiver. If the source is in motion towards the receiver, we would obtain again a negative frequency, and the result that the disturbance reaching the receiver at a given time would have been emitted by the source at a later time. The phenomena occurring when the speed of the source exceeds the speed of propagation of the emitted disturbance have been investigated by Mach;⁴ his results will not be discussed here.

The author wishes to thank Dr. G. H. Carragan of the Rensselaer Polytechnic Institute for helpful criticism and discussion.

² Now at North American Aviation, Inglewood, California.

³ For a general discussion of the Doppler effect, see H. B. Lemon and M. Ference, *Analytical experimental physics* (Univ. of Chicago Press, 1943), p. 426; or any standard textbook on sound. For a detailed discussion of the various Doppler effects, see J. P. Perrine, *Am. J. Physics* 12, 23 (1944).

⁴ For a more general discussion, see H. Bateman, *J. Acous. Soc.* 2, 468 (1931); R. W. Young, *J. Acous. Soc.* 6, 112 (1936); L. Fleischmann, *J. Acous. Soc.* 15, 103 (1943); G. F. Herrenden-Harker, *Am. J. Physics* 12, 175 (1944).

⁵ Obviously phenomena occurring at a single point would not be sufficient to define a wave propagation.

⁶ This result was already assumed in establishing Eq. (4). The assumption is based on the fact that c is a property of the medium through which the phenomenon is propagated; and, since there is no relative motion of the medium and the receiver, no change in the speed of propagation can be expected.

⁷ E. Mach, *Sitzber. Akad. Wiss. Wien. IIa*, 95, 164 (1887); 98, 1310 (1889); 105, 605 (1896).

RECENT MEETINGS

Proceedings of the American Association of Physics Teachers

The Minneapolis Meeting, June 1947

The American Association of Physics Teachers met jointly with the American Society for Engineering Education at Minneapolis, Minnesota, June 18 to 21, 1947. Paul Kirkpatrick, J. W. Buchta, C. E. Bennett and H. A. Armsby presided at the various sessions. The program was arranged by J. W. Buchta and C. E. Bennett.

Two afternoon sessions and two luncheon meetings were devoted to the following invited and contributed papers.

Invited Papers

Secondary school physics in its relation to the liberal arts college. DUANE ROLLER, *Wabash College*.

Cooperation between electrical engineers and physicists. R. D. BENNETT, *Naval Ordnance Laboratory*.

Utilization of German war research by the United States: nuclear physics; high frequency and infra-red developments. BURGOYNE L. GRIFFING and EARL A. UNDERHILL, *Wright Field*.

Engineering physics in the freshman year. HENRY HARTIG, *University of Minnesota*.

Biophysics becomes a science. OTTO H. SCHMITT, *University of Minnesota*.

Contributed Papers

1. Our future supply of scientists. PHILIP N. POWERS, *Staff Member, President's Scientific Research Board*.
2. Physicists in War Department installations. MARSH W. WHITE, *Pennsylvania State College*.
3. Practical laboratory examinations. C. N. WALL, *University of Minnesota*.
4. Demonstrations of the use of microwaves in teaching physical optics. C. L. ANDREWS, *General Electric Company*.
5. Discovering the torsion pendulum expression in the freshman laboratory. SEVILLE CHAPMAN, *Stanford University*.
6. Cathode-ray pictures in three dimensions. OTTO H. SCHMITT, *University of Minnesota*.

Many members inspected the research activities of the department of physics, University of Minnesota, and also attended various general sessions of the American Society for Engineering Education.

Oregon Section

The Oregon Section of the American Association of Physics Teachers met May 10, 1947, at Lewis and Clark College, Portland, Oregon. Fifty-six persons were present. The following program was presented:

Recent developments in luminescent materials. R. T. ELLICKSON, *Reed College*.

Oscilloscope sweep circuit. M. A. STARR, *University of Oregon*.

"The cold front" and "The warm front," Navy training films. J. A. DAY, *Oregon State College*.

Microwave oscillator demonstration. D. E. ATKINSON, *Salem*.

Eye comfort. W. WENIGER, *Oregon State College*.

Measurement of ultrashort time intervals. S. H. NEDDEMAYER, *University of Washington*.

Sixty-two years of teaching physics. W. B. ANDERSON, *Oregon State College*.

The adoption of the value 10^{-7} for μ_0 and the definition of the "candle" in terms of a blackbody radiator. W. R. VARNER, *Oregon State College*.

Tour of the Lewis and Clark College campus. A. A. GROENING, *Lewis and Clark College*.

As the result of a study made by a committee of the Section, it was voted unanimously not to form a physics section of the Oregon Academy of Science.

R. T. ELLICKSON, *Reed College*, was elected *President* and W. R. VARNER, *Oregon State College*, *Secretary*, for 1947-48.

WILLIAM R. VARNER, *Secretary*

District of Columbia and Environs Section

The eighth annual meeting of the District of Columbia and Environs Section of the American Association of Physics Teachers was held at the United States Naval Academy, Annapolis, Maryland, on May 17, 1947. The group was welcomed by Capt. W. L. Field, USN, Head of the Department of Electrical Engineering, U. S. Naval Academy. Seventy-two members and guests attended. The following papers were presented at the morning session.

The presentation of inertia. VOLA P. BARTON, *Goucher College*.

Rocket power. R. B. KENNARD, *Naval Ordnance Laboratory*.

On the phenomenological approach to fluid dynamics. R. J. SEEGER, *Naval Ordnance Laboratory*.

Evaluating student opinion of physics courses. G. SCHWARTZ, *Johns Hopkins University*.

Interferometer methods applied to two-dimensional flows. W. M. COATES, *Post Graduate School, U. S. Naval Academy*.

Infra-red signaling. W. P. CUNNINGHAM, *Post Graduate School, U. S. Naval Academy*.

Microwave demonstrations. T. B. BROWN, *George Washington University*.

Three-dimensional presentation on cathode-ray tubes. J. L. DALEY, *U. S. Naval Academy*.

Delayed conductivity. R. L. FELDMAN, *Roosevelt High School, Washington*.

The transition from thin to thick lenses. E. W. THOMSON, *U. S. Naval Academy*.

Luncheon was served at the North Severn Officers' Club, after which the group spent an hour inspecting the laboratories of the Department of Electrical Engineering at the Naval Academy. The afternoon session consisted of a presentation of the following experiments developed for demonstration lectures in the general physics course at the Academy.

Sample of lecture charts: thermometer scales. E. W. THOMSON

Momentum: (a) gun ballistic pendulum, zero vector and algebraic momentum; (b) billiard balls on wires. H. E. CARR.

Resonant (singing) tubes. E. R. PINKSTON.

Chladni plates by magnetostriction. E. R. PINKSTON.

Cryophorus, dry ice with CO_2 , freezing by reducing pressure. R. A. GOODWIN.

Adiabatic expansion, formation of clouds. E. R. PINKSTON.

Ripple tank using intermittent air jet. H. E. CARR and W. CONNOLLY.

Geiger counter, radioactivity, using 20-in. lecture-room oscilloscope to show discharge. STAFF.

Explosion by shock, yellow phosphorus and carbon disulfide. W. M. SMEDLEY.

At the business session the following officers were elected: E. W. THOMSON, U. S. Naval Academy, *President*; G. M. KOEHL, George Washington University, *Secretary*; T. B. BROWN, R. L. FELDMAN and R. MORGAN, *Executive Committee*.

GEORGE M. KOEHL, *Secretary*

University of Iowa Colloquium for College Physicists

The annual University of Iowa Colloquium for College Physicists was held this year on June 12 to 14, 1947. Those registered included 115 physicists, from 69 institutions in 18 states, and 33 staff members and graduate students from the University of Iowa. The program follows.

Glimpses of nuclear physics in 1947. L. A. TURNER, *State University of Iowa*.

New high voltage accelerators. W. W. SALISBURY, *Collins Radio Company*.

A physicist advises General MacArthur. G. W. FOX, *Iowa State College*.

Demonstrations and exhibits. W. D. BEMMELS, *Ottawa University*; M. B. BRENNEMAN, *Davenport, Iowa*; D. L. EATON, *State Teachers College, De Kalb, Illinois*; F. E. CHRISTENSON and J. W. BUCHTA, *University of Minnesota*; R. T. HARLING, *Albion College*; Z. V. HARVALIK, *Rolla School of Mines*; H. JENSEN, *Lake Forest College*; M. OLSON, *State Teachers College, Milwaukee, Wisconsin*; R. R. PALMER, *Beloit College*; A. G. ROUSE, *St. Louis University*; H. K. SCHILLING and I.

RUDNICK, *Pennsylvania State College*; C. R. SMITH, *Aurora College*; M. N. STATES, *Central Scientific Company*; L. A. TURNER, *University of Iowa*; J. A. VAN DEN AKKER, *Institute of Paper Chemistry*.

Electronics and optics. JOHN W. FORREST, *Bausch & Lomb Optical Company*.

Certain aspects of cosmic rays. MARCEL SCHEIN, *University of Chicago*.

The role of nuclear physics in undergraduate instruction. DUANE ROLLER, *Wabash College*.

The social significance of science teaching. WENDELL JOHNSON, *State University of Iowa*.

Experience with physics courses in general education. Round table discussion. W. P. GILBERT, *Lawrence College*; J. W. HORNBECK, *Kalamazoo College*; R. R. PALMER, *Beloit College*; DUANE ROLLER, *Wabash College*; IRA FREEMAN, *Swarthmore College*; C. L. HENSHAW, *Colgate University*; G. W. STEWART, *University of Iowa*.

Personalities in the Manhattan atomic bomb project. J. W. KENNEDY, *Washington University*.

Survey of low temperature physics. SAM LEGVOLD, *Iowa State College*.

New short-wave electronic tubes. J. J. LIVINGOOD, *Collins Radio Company*.

Other features of the Colloquium were an exhibit of recent physics motion pictures, several luncheon and dinner sessions, and a reception for members and guests at the residence of Professor G. W. Stewart.

One unusual feature, not on the program as arranged by Professor Stewart, was the presentation to him of a bound volume of more than 100 personal letters of appreciation from participants in past annual meetings. The presentation was made in the name of the group by H. K. Schilling, in the absence of J. C. Stearns, who had originally been selected by the participants to conduct the ceremony.

AAAS Registration Fee for Annual Meetings

As a member society of the American Association for the Advancement of Science, the American Association of Physics Teachers bears its share of concern for the financing of large general meetings of the AAAS in which we participate. The estimated deficit for the AAAS meeting in Chicago, December 1947, is \$14,000. The AAAS administration has opened its books to the AAPT Executive Committee and discussed very fully the possible sources of income which might be used to cover the expected deficit.

After thorough study the AAAS has proposed that its constituent societies restrict attendance at their sessions to persons who have paid the AAAS registration fee of \$2.00

for AAAS members or \$3.00 for nonmembers. Apparently this plan is being accepted by the great majority of the societies planning to meet at Chicago, and it has been approved by the AAPT Executive Committee, not with enthusiasm but as the only available method of carrying out our responsibilities toward the AAAS.

Persons planning to attend the annual meeting of the AAPT at the University of Chicago, December 29-31, 1947, are hereby advised that the charge will be in effect. Nonmembers of AAAS will be given the alternative of registering free if they join the AAAS at the Chicago meeting.—PAUL KIRKPATRICK, *President*.

NECROLOGY

Professor G. E. M. Jauncey, 1888-1947

Washington University suffered a great loss in the death of GEORGE ERIC MACDONNELL JAUNCEY on May 19, 1947. He had been with the University continuously since his first appointment as an instructor in 1920. Born in Adelaide, South Australia on September 21, 1888, he received his bachelor's degree at the University of Adelaide. He then went to Leeds, England, to do graduate work under SIR WILLIAM BRAGG, who had recently moved from Adelaide to Leeds. He held instructorships in succession at Toronto, Lehigh, Missouri and Iowa State.

On joining the faculty at Washington University in 1920, JAUNCEY began work on x-rays and rapidly attained fame as one of the leading experimentalists in the field. His particular area of interest was the study of x-ray scattering, especially diffuse scattering. Between 1920 and 1940 he published over 90 papers in various physics journals. In addition, the publication of many papers by his graduate students owes much to his stimulating supervision and penetrating insight. Though he was primarily an experimental physicist, he had, to an unusual degree, an insight into theoretical physics and an ability to use it to supplement his experimental work.

JAUNCEY had always been interested in the history of physics and the development of physical ideas. As long as he was able to carry out experimental work, his interest in the history of physics necessarily took second place in his activities. A few years ago his health made it necessary for him to abandon experimental work, and he turned with his usual enthusiasm and concentration to the history of physics. He planned a series of papers, eventually to be brought together in a book, on the birth and early history of various fields in physics, of which two have been published. JAUNCEY's interest in the history of physics did not follow the classical scientist's approach of cataloging every discovery and its date. He was more interested in the human aspects involved. He studied the circumstances that led a scientist to a brilliant and radical proposal. Correspondingly, he puzzled over the reasons for the failure of earlier workers to make the same discoveries. To any one who charges scientists with cold-bloodedness and an inhuman approach to life, JAUNCEY will stand as the outstanding example of a scientist who outdid the

humanists in their own field. The tremendous human interest and understanding displayed by JAUNCEY was entirely free from sentimentality and was strictly under scientific control.

In addition to his activity with the field of x-rays, JAUNCEY was keenly interested in the teaching of physics and in the perplexing problem of deciding what is fundamental and what can be bypassed in arranging curriculums in this day when the subject matter is accumulating so rapidly. Because he prepared his lectures with the utmost care, the courses he taught were regarded as outstandingly good. A third edition of his well-known and widely used book, *Modern Physics*, is now in the press. With A. S. LANGSDORF he published in 1940 a small book on the mks system of units, *M.K.S. Units and Dimensions and a Proposed M.K.O.S. System*.

Seven years ago JAUNCEY had a severe illness from which he never regained his previous level of health; yet he continued to work as regularly and methodically as he had done before his illness. His more intimate friends, who were the only ones who knew more than the barest details of his poor health, felt that his objective approach to his own misfortunes transcended by magnitudes the classical example of Socrates and the hemlock cup. He kept a notebook record of his own case, and when on some days he felt particularly ill, he would turn back several weeks and note that he was now really very much better than he had been then. His body literally wore out, long before it should have, and while he obviously knew it, he worked and planned with virtually utter disregard of it. The physical limitations which his health imposed upon him were arranged so as to be very nearly unnoticeable to his colleagues and students. His active mind, however, had laid out plans for papers and researches scheduled for completion as far as two years in the future. His human interest and understanding patently did not include self-pity. His friends and colleagues will always remember the courage and good humor with which he carried on to the day he died.

A. L. HUGHES
R. N. VARNEY



DIGEST OF PERIODICAL LITERATURE

Demonstration Potentiometer

For many purposes time is saved and maintenance simplified if students who are primarily interested in the potentiometer as a tool are provided with a complete unit ready for use, including batteries, rheostats, switches and the necessary wiring. This method has the disadvantage of obscuring the operation of the device. To remedy this difficulty, an additional unit is constructed, with all parts in the same places as in the ordinary one, but assembled in a case of transparent Lucite. The wiring in the demonstration unit is arranged to facilitate tracing the circuit. In both cases the galvanometer is a separate unit. P. BENDER, *J. Chem. Ed.* 24, 195 (1947).

Fundamental Experiments in Electromagnetism

A long Robison magnet is supported vertically on the pan of a balance and counterbalanced. The upper end of the magnet projects into a circular coil of wire wound on a suitable form. A current is set up in the coil, and the resulting force on the magnet is measured by changing the counterbalancing weights. By using coils of different diameters it is shown that the magnetic field at the center of a circular coil is inversely proportional to the radius of the coil. The current is kept the same in successive coils with the aid of an ammeter, but its readings are not used in verifying the aforementioned relation.

To verify the equation $\epsilon = -d(n\phi)/dt$ for induced emf, the emf, induced when a Robison magnet is pushed into a solenoid, is balanced against the potential drop appearing at the terminals of a resistor of known low resistance—say 0.1 ohm—that contains current. The potential drop is measured with a voltmeter, or is calculated from the known resistance and current. The solenoid is connected to the resistor by way of a galvanometer, and the magnet is pushed into the solenoid with such speed that the deflection of the galvanometer is zero. Since the number of turns of the solenoid is known and the speed of the magnet can be measured, the induced emf is readily calculated. H. F. BOULIND, *Sch. Sci. Rev.* 28, 329–39 (1947).

Demonstration of Lissajous Figures

The Pfandl apparatus for demonstrating Lissajous figures—slotted disks vibrating on springs—may be made to act continuously if the disks are mounted on one end of flat strips of metal or wood that are pivoted near their other ends and provided with short longitudinal slots. In these slots work two pins mounted eccentrically, one on each of two triple pulleys. One pulley has diameters in the ratio 1:2:4; and the other, 1:3:5. The pulleys are driven by strings passing around a single large pulley with two grooves of equal diameters and provided with a crank handle. The strings are kept taut by idler pulleys held by

coil springs. Nine combinations of ratios of frequency of vibration are thus provided.

A small lamp behind a frosted glass produces a spot of light that is visible through the crossed slits and that moves as the slits oscillate, tracing out the figures. Persistence of vision is sufficient to make the figures seem continuous. The device is easily adapted for projection.

The diameters of the triple pulleys must be very accurately in the ratios specified. If allowance is not made for the radius of the string to be used, all the figures except that for the 1:1 ratio will display a slow drift of phase difference. Any desired phase difference may be introduced by slipping one pulley to the appropriate setting of its eccentric pin. A. D. BULMAN, *Sch. Sci. Rev.* 28, 343–6 (1947).

Precipitation in the Atmosphere

An early theory of the formation of precipitation in the atmosphere was that humid masses of air of different temperatures were mixed; the mixture was then too cold to hold all the moisture present, and some of it condensed out. However, this action would occur on too small a scale to account for even a light rainfall.

The present theory is that of dynamic cooling in rising air currents. Air can be lifted by (i) convection; (ii) forced ascent of an air current over a cold wedge; (iii) lifting of warm air by undercutting cold air; (iv) orographic lifting. In addition to the cooling resulting from such lifting, air may be cooled by radiation to space and by contact with a cold surface.

Atmospheric condensation can occur only on nuclei, which are always abundantly present, mostly as fine salt particles. Water vapor condensed as a cloud can be regarded as a colloidal suspension of water in air. It may be stable or unstable; in the latter case only is there coagulation of the droplets and consequent precipitation. Stability requires (i) uniform electric charge on all droplets; (ii) uniform size of droplets; (iii) uniform temperature in the cloud; and (iv) uniform motion of parts of the cloud.

Raindrops result principally from coalescence of droplets; it has been found that the masses of raindrops are integral multiples of a standard mass. A recent theory of the formation of snow is based on the fact that at temperatures below freezing, the saturation vapor pressure with respect to ice is less than with respect to undercooled water. Liquid water droplets may exist at temperatures as low as -20°C . If ice crystals are introduced into a cloud of such droplets, the crystals will grow at the expense of the droplets by virtue of the lower vapor pressure over ice. R. J. BOUCHER, *J. Chem. Ed.* 24, 204–6 (1947).

Refresher Courses for Secondary School Teachers

Universities should offer a course on recent developments in science designed to keep the secondary school science teacher up to date. If the course is given in summer sessions, the content should be changed periodically, so that a teacher could return every five years, say, to be brought abreast of current changes. It could be given in two sections, one for the physical sciences and one for the

biological. It should be on the graduate level, perhaps conducted as a seminar. The emphasis should be on basic principles, applications and the social significance of new developments; trips to laboratories might well be included. It should be of great benefit in helping the busy secondary school teacher to keep in touch with recent advances so that he can speak with authority about them in his classes. P. N. POWERS and W. H. STICKLER, *Sch. Sci. Math.* **46**, 811-7 (1946).

New Members of the Association

The following persons have been made *members* or *junior members* (*J*) of the American Association of Physics Teachers since the publication of the preceding list [*Am. J. Physics* **15**, 360 (1947)].

Baker, D. K. (*J*), University of Pennsylvania, Philadelphia, Pa.
 Bate, George L. (*J*), Wheaton College, Wheaton, Ill.
 Bolling, N. F., Arkansas State Teachers College, Normal Station, Conway, Ark.
 Halton, Father Edward B., St. Mary of the Springs College, Columbus 8, Ohio.
 Hanau, Richard, University of Kentucky, Lexington, Ky.
 Hartley, W. H. (*J*), 4037 Neilson St., Philadelphia 24, Pa.
 Isenberg, I. (*J*), University of Pennsylvania, Philadelphia, Pa.
 Johnstone, John H. L., Dalhousie University, Halifax, N. S.
 Paton, Mary Ann, 1005 S. Busey Ave., Urbana, Ill.
 Peirce, Robert V. (*J*), 468 E. Market St., Germantown, Ohio.

Rhodes, J. L., 3723 Spring Garden St., Philadelphia 4, Pa.
 Richards, James A., Jr., R. D. 1, Olivet, Mich.
 Schmitt, Roland W. (*J*), 102 W. 16 St., Austin, Tex.
 Shank, Charles H., 744 Village Rd., York, Pa.
 Shera, Jane Ann (*J*), 111 East Walnut St., Oxford, Ohio.
 Sweeney, Joseph P. (*J*), 1304 Burnett St., Wichita Falls, Tex.
 Voelker, C. H., Physics Laboratory, American Air Filter Co., 215 Central Ave., Louisville, Ky.
 Wahlig, C. F., 101-11 115 St., Richmond Hill, N. Y.
 Warschauer, Douglas M. (*J*), New York University, University Heights, New York 53, N. Y.

Wherever is gathered together a staff of teachers competent to give training and instruction of a high quality in the college and university grades, there will inevitably appear an activity in research and production—be it in science, or in the professions, or in the fine arts. Such activity is one of the normal instincts of any person who is a live student of his subject. It not only makes him a better teacher, but also creates an atmosphere and supplies ideas and problems which have an important role in the development of his pupils.—K. T. COMPTON.